Open Research Online

The Open University's repository of research publications
and other research outputs

# The virtual participant : story telling in a computer supported collaborative learning environment

## Thesis

How to cite:

Masterton, Simon J. (2001). The virtual participant : story telling in a computer supported collaborative learning environment. PhD thesis The Open University.

For guidance on citations see FAQs.

© 1999 The Author


https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.21954/ou.ro.0000d49d

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

31 0252141 1

UNRESTRICTED

# The Virtual Participant: Story telling in a computer supported collaborative learning environment

by

Simon J. Masterton

BSc Honours, Physics, University College London
(1994)

MSc, Artificial Intelligence: Knowledge Based Systems, University of Edinburgh
(1995)

This thesis has been submitted for the degree of
Doctor of Philosophy in Cognitive Science

Discipline: Knowledge Media Institute

at the

Open University

September 1999

AUTHOR NO : M7206757
DATE OF SUBMISSION : 17 SEPTEMBER 1999
DATE OF AWARD : 5 APRIL 2001

Paginated blank pages are scanned as found in original thesis

No information is missing

# Acknowledgements

This part of the thesis is intended to test my memory of all those who have in some way helped me get to this point.

First of all I must thank Mike Brayshaw, without whose convincing I would never have applied to do this PhD in the first place. Then I must go on to thank Marc Eisenstadt who originally fired my interest in Artificial Intelligence. Most importantly of all I must thank my supervisor Stuart Watt, no one else would have put up with me.

Next I must thank Trevor Collins and Paul Mulholland who have both provided timely advice, criticism, and companionship in the local bars. I must also thank all others in the Knowledge Media Institute who have helped me out at one time or another, particularly Hannah Walton and Ortenz Rose, without whom my questionnaires would never have been sent.

From the Creative Management Team I must specifically thank John Martin for allowing me to conduct this work. I must also thank Jane Henry & Mo Vernon and the rest of the course team. The VP would never have functioned were it not for the B882 computer tutors and I would like to take this opportunity to thank all of them for their help.

This research was sponsored by the EPSRC studentship No. 95306409 and a CASE award from British Telecom. From BT I would like to thank Hyacinth Nwana, Divine Ndumu, and Nader Azarmi of the Agents group, and Neil Finlayson of the Human Factors group.

# Abstract

This thesis presents a study of a novel approach for supporting students in text based electronic conferencing. It describes the development of a concept known as the Virtual Participant. An initial prototype was developed which was tested on the Open University Business School MBA course on Creative Management. The Virtual Participant first presented itself to the users as Uncle Bulgaria, a metaphor for collecting and recycling important information.

The Virtual Participant approach is to store the discussions students have had in previous years that the course has run, and to retrieve those discussions at a time most appropriate to helping the students studying this year. It was never intended to provide 'the answer' but rather examples of similar discussions on similar topics. Uncle Bulgaria interacted with the students over a period of 16 weeks, during which time the students prepared two assignments and completed the first half of the course. The information gained from the students' interactions with the system and their feedback to a questionnaire survey was then fed back into a second prototype which was again tested on the same course.

In the second study the system was known to the students as the Active Archive, an active component of an archive of past student discussions. Through cross year comparisons it was possible to evaluate the improvements made between the Active Archive and Uncle Bulgaria systems. The Active Archive interacted with the students on a much larger scale than Uncle Bulgaria had, but with no increased negative impact. The second study provided examples where the Active Archive stimulated discussion amongst the students and vicarious learning could be said to have taken place. Taking the lessons learned from these two studies a number of guidelines for the development of such systems have been produced and are described and discussed.

# Contents

# Table of Figures

# Table of Tables

# Chapter 1 - Setting the scene

## 1.0 Introduction

In recent years there has been an upward spiral in the cost of higher education. At the same time there has been a drive to increase access. However the only way to guarantee access is to keep costs down, and keeping costs down risks quality. An approach to guaranteeing access at low cost is offered by the Mega-Universities (Daniel, 1996) (universities with 100,000+ students, e.g. the UK Open University (OU) and the Chinese Television University) a process to which he refers to as "massification". These universities have maintained quality with reduced costs and increased access.

To achieve this, Mega-Universities have made use of original multi-media technologies in their teaching, television and radio programs, pre-recorded audio and video cassettes, and paper-based materials. These universities have always made use of new technology and have recently begun to use computer supported collaborative learning (CSCL) (Hall, Miyake & Enyedy, 1997), and text-based electronic conferencing, to assist in their teaching. These technologies have their own problems. Researchers have found that despite their benefit in traditional and distance teaching they cause an increased drain on staff time and resources, not necessarily proportional to their benefit, perceived or otherwise (Nixon & Salmon, 1995).

In this thesis I present the concept of the Virtual Participant (VP). The aim of the VP is to provide a way of helping students with their studies, reducing the load on the tutors, and maximising the benefit to the students. With the move to increase the use of on-line discussion forums for students, electronic course materials, assignment submission and marking, and demos and simulations on CD ROMs, the computer has become a vital learning tool (Thomas, Carswell, Price & Petre, 1998). These Universities are exploiting the computer not only to teach, but also to guide and assist. Students are assisted in their course choice (Scott & Phillips, 1998), or when studying they may have their own personal learning manager (Sumner & Taylor, 1997), guiding them through the materials and course work.

For the Open University one key reason for high completion rates by distance students is that the courses and assignments have strict timetables and deadlines. Systems like a personal learning manager can be linked to the timetable to help students keep on course. I propose that these approaches could be combined to create

the concept of the Virtual Teaching Assistant (VTA), with the intention of supporting students on a number of levels. Students could be advised on which courses to choose, aided in managing their study of those courses, and supported in their learning from the course. The Virtual Participant provides but one component of the Virtual Teaching Assistant concept, which will be discussed in more detail in section 8.6 of chapter 8.

The thesis looks in detail at the development the Virtual Participant which is aimed at supporting students in electronic conferencing. The VP is an educational, or 'pedagogical', agent that helps to answer common student questions in an electronic conference. This thesis describes the two generations of prototype I went through in developing this system, the students' experiences with it, the lessons I have learned from them, and the future developments and applications of this system.

## 1.1 Motivation

With the massification of higher education it has been argued that the "need to support individual learners through learning dialogues becomes more pressing" (Mayes & Neilson, 1995). Mayes & Neilson present the "learning cycle", Figure 1.1, as the stages which lead to the growth of understanding. At each stage in the cycle different types of courseware exist to support the learner. Mayes & Neilson argue that "most learning technology has been directed towards the presentation and content, rather than the answering of questions, or the opportunity for discussion and reflection".

Mayes has developed his Learning Cycle idea further (Mayes, 1995), building it on the TOTE (Test-Operate-Test-Exit) unit of Miller et al. (1960). Focusing on the Dialogue stage of his cycle he expands it with a sub-cycle of Discussion-Reflection-Reification[1], using John Dewey's discussions of learning to show that the need to support reflection is recognised.

[1] Reification is the term used to indicate a stage of reconceptualisation arising from the Discussion and Reflection.

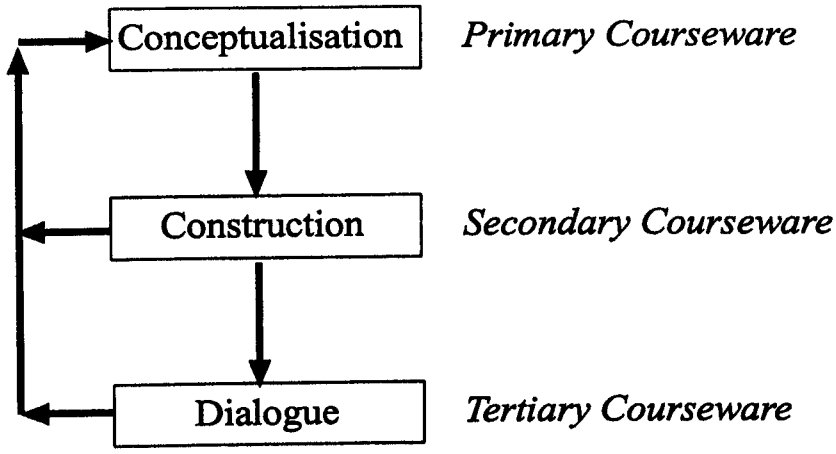


Figure 1.1. A description of learning as a cycle of stages leading to the growth of understanding (Mayes & Neilson, 1995).

Mayes' work builds on that of Kolb (1984) and Pask (1975) in emphasising the importance of dialogue in learning, although he does not cite them directly. Kolb (1984) compares the learning cycles of Dewey, Lewin, and Piaget and draws the conclusions that:

- Learning is best conceived as a process, not in terms of outcomes.
- Learning is a continuous process grounded in experience.
- The process of learning requires the resolution of conflicts between dialectically opposed modes of adaptation to the world.
- Learning is an holistic process of adaptation to the world.
- Learning involves transactions between the person and the environment.
- Learning is the process of creating knowledge.

The experiential learning cycle of Kolb can be found in Figure 1.2. It is clear that Kolb also sees the step of Reflective Observation as being quite important, parallel to the Dialogue step of Mayes' learning cycle, which also includes reflection.

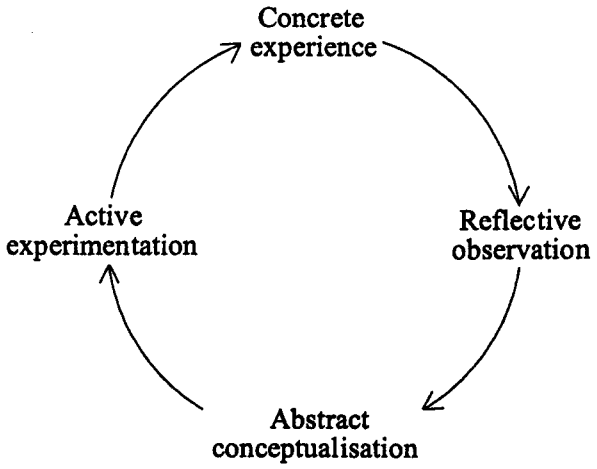


Figure 1.2. Kolbs' experiential Learning Cycle (Kolb 1984, pages 33 & 42).

Earlier still, Pask developed a 'conversational' model of learning, a simplified version of which is presented in Figure 1.3. In essence Pask's idea is that the teacher uses an entailment structure of the relationship between domain concepts from their own point of view, which is then used to guide the conversation between tutor and student. This conversation involves a continual cycle of teaching and teach-back as the student learns. In this model the student and teacher are assumed, and aided, to reflect on the learning process. Whalley (1995) takes Pask's model and identifies, for his own needs, where information technologies can support the presentation of the teacher's model to the student and vice versa ($IT^1$ and $IT^2$ in figure 1.3). Whalley then goes on to note that one other area where information technologies can help is "to further discussion between students and with their teachers - supporting the 'learning conversations' in Pask's terms". Here "$IT^3$ covers what is now known as 'computer supported cooperative learning' (CSCL)", exactly the technology the Virtual Participant is aimed at.

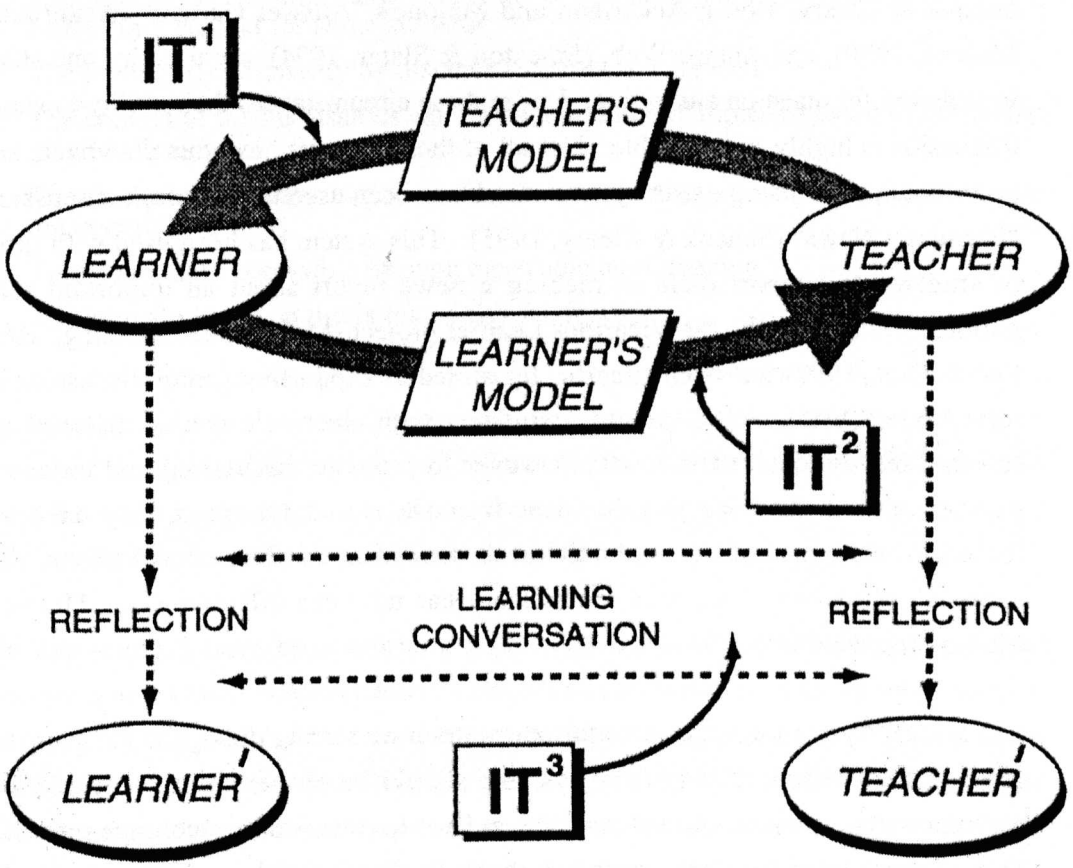


Figure 1.3. Pask's model of the shared presentation of knowledge (reproduced from Whalley, 1995).

In his learning cycle Mayes has associated different types of courseware with the three stages. Technology for content presentation, a common development in recent years e.g. (Eisenstadt, et al. 1996), is referred to by Mayes and Neilson as primary courseware. They define secondary courseware as that which supports the learner in exploring arguments and problems, e.g. (Thomas et al., 1998). The third level is tertiary courseware, to help support question answering, discussion and reflection. The Virtual Participant is intended to fulfil this role by supporting dialogue, something which the above work has identified as a vital part of learning.

A number of systems have already been developed which fulfil the role of tertiary courseware. These systems store dialogues in a structured form which is then used to support question answering, discussion and reflection. However some of these systems, for instance some of Schank's "Engines for Education" (e.g. Transcom,

Schank & Cleary, 1995), Ackerman and Malone's "Answer Garden" (Ackerman & Malone, 1990), and AnswerWeb, (Smeaton & Slater, 1994), are used by one student in isolation for question answering. Under these circumstances their ability to support discussion is highly questionable. Not all of these systems have this drawback, and a good example of a single user system which has been used to encourage discussion is "Broadcast News" (Schank & Cleary, 1995). This system has been used with groups of students to support them in making a news report about an important recent political event. Finally, the Vicarious Learner project (McKendree, Stenning, Mayes, Lee & Cox, 1997) has been specifically aimed at supporting group discussion and reflection. To this end it provides students with electronic course material with relevant discussions linked to it. However to promote discussion, and reduce the number of students using just the ideas from the stored dialogues, they have been forced to make part of the student's mark dependent on their contributions. This approach of making discussion compulsory has not been effective (Lee, Dineen & McKendree, 1997).

The tertiary systems discussed so far concentrate on storing dialogues for answering questions. I believe that tertiary systems should be concentrating on *facilitating* dialogues and only through that can they answer questions and encourage reflection. The VP embodies the facilitating approach to dialogues through the medium of electronic conferencing. Electronic conferencing has proven very popular with distance students. Conferencing makes distant tutors and peers more easily accessible and increases the feeling of community, reducing drop out rates. In this context the dialogue is facilitated by the presence of tutors in the environment who are able to stimulate discussion. However this is an intensive task requiring a lot of time and effort. One approach to stimulating discussion is to identify captured dialogues from previous years about subjects where are currently under debate. By retrieving these discussions, at the appropriate time, and presenting them to the students it is possible to widen the debate through the alternative viewpoints espoused by previous students. The VP does this by monitoring the current discussions and contextualising previous stored discussions to issues as they arise. This should reduce the burden on the tutor to stimulate discussion by presenting alternate viewpoints, and help to maintain the content level of the conferences.

## 1.2 Review of the problem

Electronic conferencing provides many advantages to the student, but at the same time these may be disadvantages to the tutors. To summarise the most significant of these points:

- The number of students using electronic conferencing is growing, and there may be no other contact for some students.
- The feeling of community by distance students is enhanced.
- The content of the discussions can be poor and not all topics relevant to the course may be discussed.
- The students may have access to discussions from previous years, but these are unstructured.
- Tutors are more accessible through electronic conferencing.
- The increased load on tutors may delay feedback.

Electronic conferencing provides a good support environment for discussion. Increased access to tutors and peers enables question asking, and the permanent record of the discussion can help reflection. On the down side, conferences without direction can rapidly deteriorate, tutors are not available all the time, and past years' discussions may be difficult and time consuming to use.

On the course I have been working with (The Open University Business School Course: B882 'Creative Management') electronic conferencing is an optional activity. Because of this the students need to see a clear benefit from conferencing to be convinced to use it. Conferencing is viewed as an important part of the educational experience and therefore how students become socially acclimatised is also important. Nixon & Salmon (1995) have developed a 5 stage model of student social acclimatisation, table 1.1. They have identified the first two stages as the most important time for the student and hence where learner support should be concentrated. The VP approach takes this into consideration as an important point in developing on-line interactions.

| Stage | Process |
|---|---|
| 1. Access | Getting past the hardware, software, network problems to getting on-line |
| 2. Induction & Socialisation | Getting to know the conferences, posting first messages, meeting people |
| 3. Seeking information | Asking questions, finding information from others |
| 4. Interaction | Helping others, participating, answering questions |
| 5. Boundary Shifting | In control, responsible for own on-line learning, student is now an "independent learner" |

Table 1.1: The 5 stages of social induction for new students using text-based electronic conferencing. From (Nixon & Salmon, 1995).

It is not clear that all students pass through these stages, or that they are all willing to progress through these stages, or even that these stages exist in the linear fashion

presented. In an electronic environment it is much easier for students to exist as 'lurkers' browsing information from the discussions without ever sending any messages. For example it would be interesting to compare this with learners in a normal tutorial group and their progression towards being "independent learners". However this broad brush approach is sufficient for our purpose of looking at new learners in general.

To summarise, conferences from previous years provide us with cases representing discussions on aspects of the course where there may be multiple points of view. These cases are not structured and hence prove a difficult resource to exploit. The Virtual Participant is a tertiary courseware system which provides access to this resource by contextualising relevant previous dialogues to current discussions, in a text based electronic conferencing system.

## 1.3 Goals of the Virtual Participant

As I have already noted massification is on the increase, and the use of electronic conferencing to allow students to communicate with each other and with their tutors is also growing. However this is at an increased cost in comparison to the economies of scale usually reaped by distance education. What is needed is some way to help the students help themselves. It has already been identified that discussion and reflection are vital parts of learning and that electronic conferencing could in principle help support discussion and reflection between students who would otherwise be isolated. However there are still some shortcomings of electronic conferencing, such as:

- Participating in electronic conferencing is a drain on tutors' time.
- Not all students like to take part in discussions, often up to two thirds will lurk.

The VP will focus both on enhancing the positive aspects of conferencing and on addressing the shortcomings. The goal of this study is to demonstrate the following:

- **The required effort of tutors can be reduced.** The VP will aim to provide content input relevant to the context of discussion - reducing the need for tutor involvement.
- **Discussion can be stimulated.** Discussion is known to be an important element of learning and because of this the VP will aim to stimulate more discussion by its input into existing discussions and through posing questions. This will improve the value of the electronic conferencing and is linked to the next point.
- **Similar discussions take place each year.** With similar course materials and assignments each year it is likely that similar discussions will take place. This study will not only show that this happens, but also that previous years discussions

can prove an important learning resource improving the experience for current students, especially when contextualised to their discussion.

- **It is possible to help students, even if they don't participate.** Not all students like to take part in discussions, often up to two thirds will lurk. The VP will aim to provide more quality content for lurkers and raise the quality level of discussions to encourage lurkers to participate. The VP will also interact with them directly, encouraging them to contribute.

The VP supports students' discussions by reflection and question answering, not directly, but by presenting previous discussions in the form of (what I am calling) a story, to encourage a debate and to allow students to reflect on the discussion. These stories come from the discussions of previous years and are structured in such a way that the VP's first contribution to a discussion is about similar problems experienced by other students. This message can be responded to by a student selecting from a choice of questions. The VP then responds by providing more details, or details about a related case. I will come back to the structure of stories in section 3.2.4, and an example interaction between the VP and a student can be found in chapter 4.

I argue that the VP's approach of contextualising previous years' discussions, these stories mentioned above, into the current discussion improves the learning experience for the students. In other tertiary courseware systems stored dialogues have been used as a substitute for real discussions, or are contextualised by linking them to relevant sections of course material. The clear difference between these approaches is that the VP is just another participant in the conference and provides an active resource rather than a passive one. In effect past discussions are contextualised with current discussions, a truly tertiary courseware approach, rather than, by augmenting course materials, a primary courseware approach.

## 1.4 Review of the approach

I have adopted an agent-based approach to the problem. This term has been much abused recently and so it may help to look at its definition. Using the Oxford English Dictionary:

- **agent, 4.** 'One who does the actual work of anything, as distinguished from the instigator or employer'.

Our agent is a type of 'interface agent' (Kay, 1990) and being a tutor's assistant, supporting student discussions, we delegate tasks to it. The teaching task is the provision to students of contextually relevant cases at times relevant to their discussions. This is like a human tutor reflecting the experiences of previous

students. In this way the current students are encouraged to draw parallels between their current situation and others being presented. The VP approach has these aims:

- Contextualising cases from previous years by presenting them in current discussions about similar problems.
- Eliminating the need for students to search for relevant cases.
- Being non-intrusive, so that its messages can be easily ignored by those not interested.
- Reducing the need for the tutor to repeatedly provide the same information.
- Providing an immediate response to the student in an asynchronous environment.

Artificial agents in distance learning are being seen as a new approach for providing computer based teaching, a framework for which is given in (Chan, 1995a). In this paper Chan presents four different roles for artificial agents: as a teacher; a companion; an assistant for the student; an assistant for the teacher. Chan's framework progresses from straightforward teaching systems through to more social learning systems, as seen in his other work (Chan, 1995b) and Schank's work on GuSS (Guided Social Simulation) (Schank & Cleary, 1995). These social learning systems involve some (possibly one) students and some (again possibly one) simulated other. Both Schank and Chan claim that the interactions between the learner(s) and the other(s) provide an overall better learning experience.

Chan's framework is limited in that he views the students' assistant as a secretary, helping in organising, prioritising and guiding. This is very similar to the approach adopted by the personal learning manager (Sumner & Taylor, 1997). The Virtual Participant, however, does not fit with Chan's framework. The VP is a teacher's assistant, however Chan's view on teacher's assistants is limited to monitoring and storing student profiles, which are then used to improve an artificial teacher's student modelling ability. The VP is intended to assist the tutor by providing what help it can to the students, and therefore extends the tutor's assistant model in Chan's framework.

Chapter 2 will develop a summary of case presentation systems related to the VP. This summary will describe how they interact with their users, the format of cases, the method of retrieval, and how they are maintained. This summary is developed through the comparison of four different kinds of case presentation systems, various aspects of which have been used as foundations for the VP. I have already mentioned two of these systems; Engines for Education (Schank & Cleary, 1995) / ASK systems (Ferguson et al., 1992) and the 'Vicarious Learner' (McKendree et al., 1997). The other two systems are FAQs (Hambridge, 1995) and Contact Finder (Krulwich & Burkey, 1996). The VP takes a similar pedagogical approach to the Vicarious

Learner, allowing students to learn by observing the discussions of others. However the VP differs from the Vicarious Learner in that it is an active system joining in with current discussions. When answering questions the VP uses a similar structure to ASK systems, presenting a similar case and discussion points around it rather than the question-answer pairs used in FAQs. The last system, Contact Finder, is similar to the VP as it suggests possible 'contacts' to participants of an electronic conference when they ask questions it can provide help with.

Finally, in educational groupware situations (e.g. conferencing and the VP) it is important to provide a clear benefit to all involved. The intention of the VP approach is it requires no extra effort with its non intrusive behaviour and benefits all by providing extra materials at relevant times. The following two sections highlight the studies which were conducted and how those studies and the timeline of the research project are reflected in the structure of the thesis.

## 1.5 The studies

The research detailed in this thesis has all been conducted with the Open University's Business School MBA elective course B882 - Creative Management. This course was the first to use electronic conferencing on the MBA profile and provided me with an archive of past discussions lasting back to 1991, ideal for a case-based approach. The course has always been at the forefront of new technological developments and is often willing to experiment with new ideas and approaches. With this in mind they agreed to try out the two experimental VP systems, Uncle Bulgaria (UB), and the Active Archive (AA).

The initial study with Uncle Bulgaria (chapters 4 and 5) helped to rapidly identify what worked and what didn't. I was able to identify important points in getting the system accepted by the students and the tutors, as well as key needs in the functioning of the system. The first study also threw up a number of key experiences in the introduction of CSCL systems such as this. A survey of a selected group of participants was conducted to get more details from their perspective.

The second study with the Active Archive enabled the trial of an improved system. This is discussed in chapters 6 and 7 along with a more comprehensive survey. Chapter 8 brings together the lessons learned from the two prototypes, combined with the experiences of related systems, to provide a guide to future developers.

## 1.6 Guide to the reader

This section is intended to guide you through what is to come in the following chapters. There are two parts to this guide. Firstly I present a time-line of the research and how that maps onto the different chapters, shown in figure 1.4. Secondly I will give a short chapter by chapter summary of what is to come to help you gain a high level view of the thesis before diving into the body of the thesis.

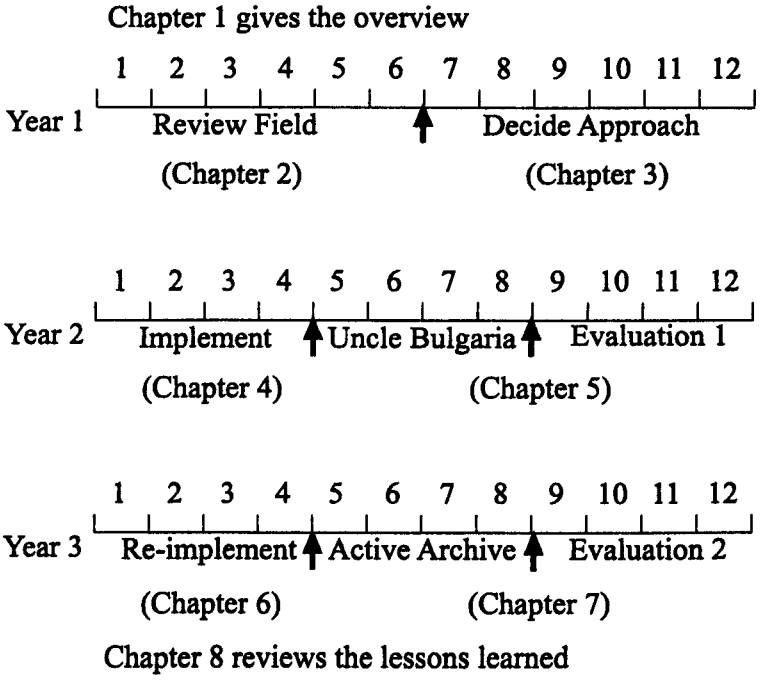


Figure 1.4. Time-line of the development of this project.

The thesis really breaks down into three sections. Chapters 2 and 3 provide an introduction and a review of the area and relevant literature with the initial work which guided our ideas. Chapters 4, 5, 6 and 7 present the two prototypes; 4 and 5 being the 1997 study; 6 and 7 being the 1998 study. Finally chapter 8 discusses what we have learned and highlight future research areas.

Chapter 2 contains a review of relevant literature. This thesis is clearly interdisciplinary spanning many areas and a full review of all possible literature which could contribute something to the discussion would be prohibitive. Instead this chapter reviews the relevant work which inspired this thesis and is intended to give a background to the ideas, motivations and reasoning. Chapter 3 is an introduction to the development of the knowledge base. This chapter covers many things including

the initial work out of which the first prototype organically grew, and the initial approach taken. It provides an overview of the techniques used in developing the knowledge base of previous years' discussions. Chapter 4 describes the development of the first prototype and gives a systems description of its functioning. Chapter 5 assesses this first prototype, its advantages and disadvantages and what the participants really thought of it. This then paves the way for our second prototype. Chapter 6 discusses the development of the second prototype taking into consideration the feedback and assessment of the first prototype and its functioning. Chapter 7 then reviews the second prototype and describes how well it fared, following a similar pattern to that of the first prototype. We collected a rather large amount of detailed data in this second evaluation. The highlights are covered in this chapter with the more detailed data in appendixes 6 through 10. Chapter 8 reviews the lessons learned from the development of the Virtual Participant, placing them into a more detailed context for those wishing to attempt similar projects, before concluding the thesis and paving the way for future work.

# Chapter 2: The Role of Artificial Agents in Distance Learning

## 2.0 Overview

This chapter draws together various different threads of research which have stimulated this thesis. To begin, we look at agents in distance education and other people's views of them, starting with Chan's (1995a) analysis of the approach and some of the roles that agents might take in the learning technologies of the future. This leads to a discussion on the possible risks involved with such approaches as an introduction to the following section on how people might interact with agents.

From here we move on to look at a different type of agent system, those that have the role of an assistant. We then discuss Don Norman's (1994) views of how people might go about interacting with agents and what problems agents might cause. Building on the points in the previous chapter about tertiary courseware and supporting dialogue, we look at these systems in general and at Answer Garden (Ackerman & Malone, 1990) and the Vicarious Learner (McKendree & Mayes, 1997) in particular detail.

Finally we present a summary of case presentation systems and review their most important features. This analysis drove much of the early implementation of the first prototype system.

## 2.1 Artificial Agents in Distance Education.

Tak-Wai Chan (1995a) presented a review and taxonomy of agents in distance education, both their possible usage and some example applications. We start with his two classes of agents in distance learning:

- "Participant agents" that assume roles in participating protocols of learning activities, e.g. the role of tutor or learning companion.
- "Assistant agents" that serve as personal assistants, intelligent tools, or surrogates on behalf of the end users in their absence, e.g. a secretary to the tutor or the student.

This research does not fall comfortably into either category. Given the amount of space dedicated in Chan's paper to the various different types of agent, he clearly views the role

of learning companion (6 pages) to be more important than that of an assistant (1 page), and the assistant's role in a multi-user environment is barely mentioned in a one sentence suggestion. From this, and the lack of any mention in other work, it seems likely that the possibilities have not been fully explored and that there is room for further work. It is the intention of this work to look at new possibilities and extensions to Chan's current definitions.

### 2.1.1 What is an agent?

The term agent has been defined and redefined by many researchers, often to fit their own work and views. The proliferation of these definitions is such that Franklin and Graesser have attempted to provide a "Taxonomy for Autonomous Agents" (Franklin & Graesser, 1996). The taxonomy developed by Franklin and Graesser can be seen in figure 2.1 below. The work detailed in this thesis is a 'Software Agent' used for educational purposes and falling between a 'Task-specific' and an 'Entertainment' agent.

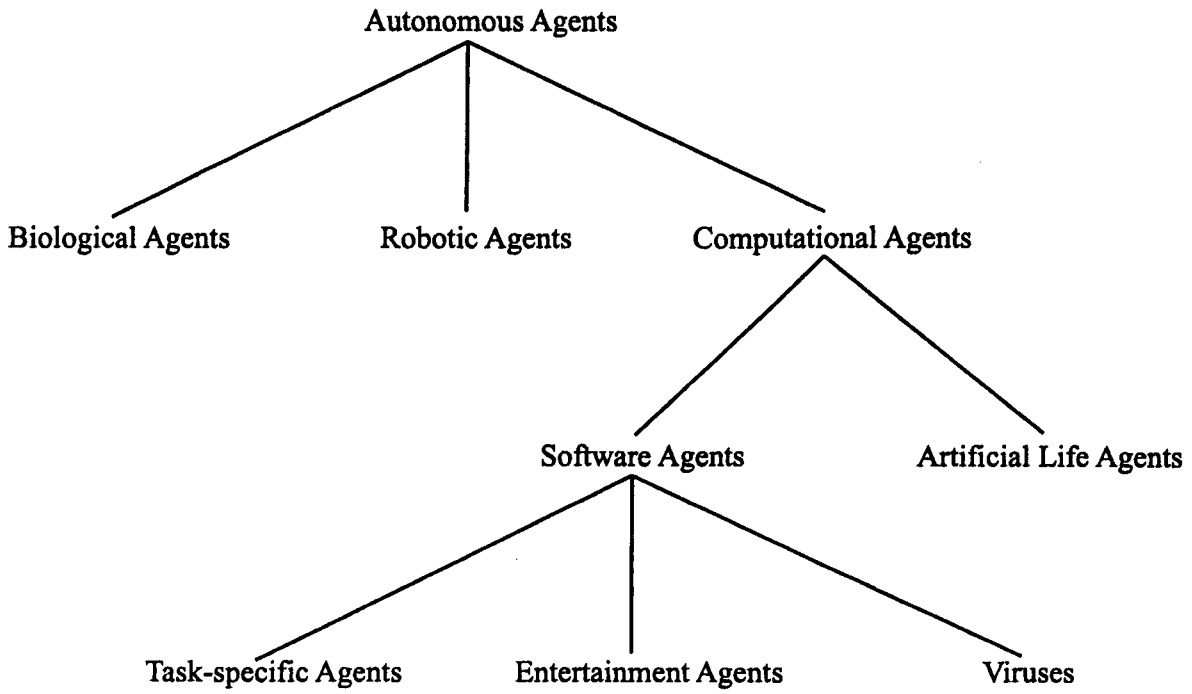


Figure 2.1: Taxonomy of Autonomous Agents from (Franklin & Graesser, 1996).

Having now partially classified our agent within an existing taxonomy let us now look at the role our agent is to fulfil. The fourth definition of agent in the Oxford English Dictionary states:

> **agent, 4.a.** Of persons: One who does the *actual work* of anything, as distinguished from the instigator or employer; hence, one who acts for another, a deputy, steward, factor, substitute, representative or emissary.

In the light of this definition the high level role of an agent is as something you can delegate tasks to, relieving yourself of the burden - performing them in a way that is acceptable. Going back to Chan's paper on agents in distance education (1995a) he also suggested two different ways of viewing an agent's capabilities, which are briefly summarised below.

*Functional View:*

This view describes the functions and services we might wish of an agent.

- Agents function continuously and autonomously in an environment with other processes and agents.
- May participate in some activities with end users; often acting on behalf of its owner.
- May assist owner in some specialised task.
- May be personified in a style which indicates its abilities.

*Operational View:*

This view provides an abstraction of the physical nature of an agent. Chan has chosen to only consider the operational view from the knowledge level (Newell, 1982), to present three different levels of operational complexity an agent may have.

- Reflex agents: Production rules acting on perceived current state of the environment.
- Agents with memory: Aware of current environmental state with stored information about previous events.

- Human-like agents: Complex agents to which we ascribe beliefs, desires and intentions. This definition is drawn from (McCarthy, 1979).

It is not clear what relationship Chan sees between the functional and operational views. He presents no guidelines, or even examples, for distinguishing or relating the views. However, this does provide an initial taxonomy of possibilities and if you were to matrix the views together it would be interesting to map out where various systems fall in terms of required intersections, possible intersections, and conflicting intersections. For example you would expect a human-like agent to be personified in a style which indicates its abilities.

### 2.1.2 Roles of artificial agents

So far we have seen a taxonomy of autonomous agents, a definition of the term 'agent', and a review of two different ways to view the functioning of agent systems. However, and rather importantly, when giving the definition of the term 'agent', the idea of roles was mentioned. Chan also sees agents in terms of the roles they can fulfil and dedicates space in his paper (1995a) to the following four: artificial teacher (1/2 page), artificial learning companion (6 pages), student's assistant (1/2 page), and tutor's assistant (1/2 page). These relative amounts of space devoted on each role may serve to indicate the relative importance Chan holds for them. The following subsections present the details of Chan's agent's roles ending with a comparison between these and the approach of the Virtual Participant.

*Artificial Teacher*

This is an idea that has been around since computers were first seen as a potential tool for teaching and learning. An agent's role in this case is to diagnose errors and give advice to the students while taking part in a special case of social learning, that of two agents interacting. Chan states that the main challenge facing the agent is ensuring the adaptability of the tutoring dialogue, which is dependent on the student model. Chan presents the agent as being similar to a traditional intelligent tutoring system, incorporating domain knowledge for both the problem space and teaching strategies. It appears that his vision for such a system is simply as a new front end to the classical Intelligent Tutoring System (ITS) approach. In taking this approach, as with any of the agent choices, there is the risk of falling into the trap of believing 'we know best'

(Eisenstadt et al., 1992) when it comes to interacting with the students. How the student perceives the role of the system is what is most important; if they feel they are being told what to do instead of being taught then this could prove problematic for their acceptance of the agent.

*Artificial Learning Companion*

The concept behind this agent role is the simulation of a companion for the student. Chan describes the learning companion approach as one where the agent "learns together and interacts with the student under the guidance of the teacher". In the majority of cases he seems to intend that the teacher is an ITS and the learning companion is simulated as part of the whole system. Chan then goes on to describe five subroles for an artificial learning companion: an artificial co-operator or collaborator, artificial competitor, artificial tutee, artificial group leader, and artificial roles in role-playing game learning.

The idea behind an artificial co-operator or collaborator is based on the theory that a teacher is not necessarily required to transmit knowledge but instead the computer can collaborate with the student on the learning of the task. An artificial competitor, on the other hand, is there to provide the student with an additional challenge, and hopefully increased motivation for the completion of the task. Chan, however, does note that the use of competition in education can be a sensitive issue and that it should be done prudently. An artificial tutee proves particularly useful in the education of teachers. In this case we are inverting the normal teaching model by having the student teach the learning companion. An artificial group leader performs a task more akin to a management role. This agent is used to keep track of groups of students working on a common project and is useful in a distributed network environment. Finally we have agents taking up artificial roles in role-playing game learning. The approach here is to use situated learning in which the student takes one role, and the agent adopts the role of another character in the environment.

*Artificial assistants*

In comparison to the above two categories Chan gives very little detail on the roles of assistants and presents quite a narrow view, perhaps because the majority of his work is on the other dimensions (e.g. Chan, 1995b). Chan simply presents assistants as a way of automating mundane tasks, without any role in the tutoring of the student. There is no

consideration of the wider roles available on an individual basis, like the personal learning manager (Sumner & Taylor, 1997), or a group basis, like the Virtual Participant.

*Artificial assistants for students*

Chan sees a student's personal assistant as desirable in many circumstances. The assistant can be used to prioritise information and messages from other students and the tutors. It is possible that tutors would require such a system to be limited in power for the student, always prioritising the tutor's messages, irrespective of the student's wishes. The assistant could take on the role of a personal secretary, booking meetings, managing the students' calendar and suggesting a schedule of study. An artificial assistant may also be useful in searching for 'other learning resources' from databases.

*Artificial assistants for tutors*

Chan presents these assistants as a way for the tutor to be able to observe the students continually, even though the students might not want it. The agent is designed to build up profiles of the students to assist the tutor in tracking their progress. Chan's intention is that the profiles generated could be reused in the teaching strategies of an artificial teacher. This would help the current students, and perhaps also future students, by identifying common problems and common types of students.

*Summary*

These agent roles provide a framework which we can use to classify many of the other pedagogical agent systems. Although some systems may fall into more than one category this still allows us to draw parallels between and inferences about different systems. The Virtual Participant is perhaps best thought of as a hybrid of personal assistant for the tutors and a learning companion for the students as the boundaries between these roles are blurred. It aims not to tutor students but rather to support them, as section 1.3 of the introduction made clear. It is important to note that the perceptions of the students, or the users of the VP in general, can affect their use of a system such as this. The next section looks specifically at how people might interact with agents, highlighting some of the other risks.

## 2.2 'How might people interact with agents?'

In (Norman, 1994), "How might people interact with agents?", Don Norman initially points out that the concept of an 'intelligent agent' generates fear, fiction and extravagant claims, seen clearly in Jaron Lanier's paper "Agents of Alienation" (Lanier, 1995), his discussion with Pattie Maes on "Intelligent Agents = Stupid Humans" (Lanier & Maes, 1996), and topically in Leonard Foner's paper "What's An Agent, Anyway? A Sociological Case Study" (Foner, 1993) What are the special problems that agents cause? Norman (1994) lists six points which he considers perhaps the most important factors to be considered:

- **Ensuring that people feel in control of their computational systems.**
- **The nature of human-agent interaction.**
- **Built-in safeguards to prevent runaway computation.**
- **Providing accurate expectations (and minimising false hopes).**
- **Privacy concerns (a subset of feeling in control).**
- **Hiding complexity while simultaneously revealing the underlying operations.**

When you think about these factors in a group conferencing environment a whole new level of complexity is added. Who is allowed to control the system? Can all users make choices which affect all others? Can individuals choose whether they see messages from an agent or not? How can you **ensure that people feel in control of their computational systems?** Should an agent's messages be filtered off to an 'observations' conference so as not to interfere with the flow of discussion. Perhaps the best way to address these points is to consider the agent as being rather like another user; if you do not like what they say then you can always choose to ignore them (given that runaway message sending is prevented). Netiquette (Hambridge, 1995) developed as a direct way of dealing with these problems including blanket technical solutions like the 'kill list'[1].

---

[1] A 'kill list' serves to help the user's mail system delete messages automatically from people the user does not wish to hear from.

**The nature of human-agent interaction** is perhaps the most problematic. Do you opt for the privacy option, allowing students to interact directly with the agent using private messages? Or do the interests of the many outweigh the needs of the one - the interactions should be public so other students can learn vicariously from them? Addressing interaction and **privacy concerns** is a key factor in whether a collaborative filtering system is successful or not (Resnick & Varian, 1997) and an issue I will come back to in the next section.

Is it possible to **build-in safeguards to prevent runaway computation?** And what are the penalties to the agent from this? Is it just an inconvenience or does it cause a kind of 'mission critical failure' where the students reject the agent outright? The solution to this is to gain the trust of your users, so that they are content with the system. Any failures would then be seen as minor aberrations and not a reason to stop using it.

One major problem with the anthropomorphism of agents and agent systems is that they may fail to **provide accurate expectations** for the users. This was a real concern for the tutors involved in the study, and it is something I will come back to again later in the thesis. Natural language understanding systems have often experienced the problem where users over- or under-estimate the system's abilities. Users then become angry or surprised when the system performs counter to their expectations. Choosing the correct level of anthropomorphism is very difficult and may well vary with different users. Some clear examples of anthropomorphism can be seen in the Maxims and NewT systems (Maes, 1994) discussed in the next section.

The final issue is to hide the agent's **complexity while simultaneously revealing** to users - be they students or tutors - **the underlying operations**. This problem is closely related to the level of trust the users hold in the system. In the development of the VP the situation arose where to gain the trust of the tutors they requested that they be given access to all controls on the system. However in doing this the level of trust reached was such that they were not actually interested in using this control. I will return to this point in chapter 6.

How people interact with agents is very complex. With agents specifically the interaction varies a lot from one context to another. It is important for future designers of such systems to learn from the past experiences of others so that these lessons do not need to be learned again, and can be adapted for each new situation. We have already tried to

move towards this with a paper reviewing some of the issues in this area (Masterton & Watt, 1999), and again I come back to this later in the thesis.

## 2.3 Agents as assistants

In Chan's view of the roles agents can play (1995a), he proposes assistants for both tutors and for students. Outside the world of education there has been a lot of other work done on the nature of agents as assistants. I have separated this work into three categories;

- **Agents that you delegate tasks to.** These agents are controlled by their users and perform specific tasks for them.
- **Agents that function autonomously.** These agents are not directly controlled by their end users, but rather provide a service.
- **Collaborative filtering systems.** These systems are not usually described as agents but their role is that of an assistant. Their groupware characteristics and problems are also particularly relevant to the VP approach.

These systems all represent different approaches and aspects of the use of agents, be it for the task of filtering your personal e-mail, joining in on a brainstorming session with related information, or helping you find interesting messages in Usenet news by collaboration. There are a number of useful lessons to be learned about these different types of system, and the uptake by their users.

*Agents that you delegate tasks to*

In this section we will describe three systems which have been designed as personal assistants with the idea of delegation in mind. Firstly we will look at two classic systems, Maxims and NewT, from the Intelligent Agents group at the MIT Media Lab (Maes, 1994). Both of these systems provided a filtering service for their users prioritising and deleting information based on profiles of their users. The first, "Maxims", filtered personal e-mail and the second, "NewT" filtered Usenet news. In both cases the systems worked in isolation, dedicated to single users. However extensions were suggested which would allow a more collaborative filtering approach with other users. The third system we describe, Luigi (Watt, 1996), is intended to take over the secretarial function of organising meetings, one of the suggested tasks for a student's assistant (Chan, 1995a).

Maxims (Metral, 1993) was a single electronic mail agent which learned to prioritise, delete, forward, sort and archive mail messages on behalf of its user. NewT (Sheth, 1994) was a news filtering system, performing the same tasks as Maxims but using multiple agents based on the users interests. In both these systems the agents had a visual presence within the user's environment and learned by example. These examples were then used to filter new messages.

Maxims learned user preferences by continually observing how the user dealt with their e-mail messages, looking for patterns and common behaviours. The user controlled the level of automation by setting two thresholds. The first, "tell me", determined the level of confidence the agent must have in a behaviour before it notified the user of something that it could do. The second, "do it", threshold determined whether the agent should directly take autonomous action on behalf of the user. A separate part of the user feedback was in the form of facial expressions which indicated the current state of the system. Figure 2.2 shows the association between threshold and feedback. In addition to the learning aspect of Maxims it was possible for the user to directly instruct the agent about what to do in a specific situation. A proposed additional mechanism for Maxims in dealing with new situations was multi-agent collaboration. By enabling agents belonging to users in a local community to interact among themselves they would be able to get pointers as to how others dealt with certain types of e-mail. The aim was to provide maximum benefit to the user for the lowest effort on their part in the fastest time.

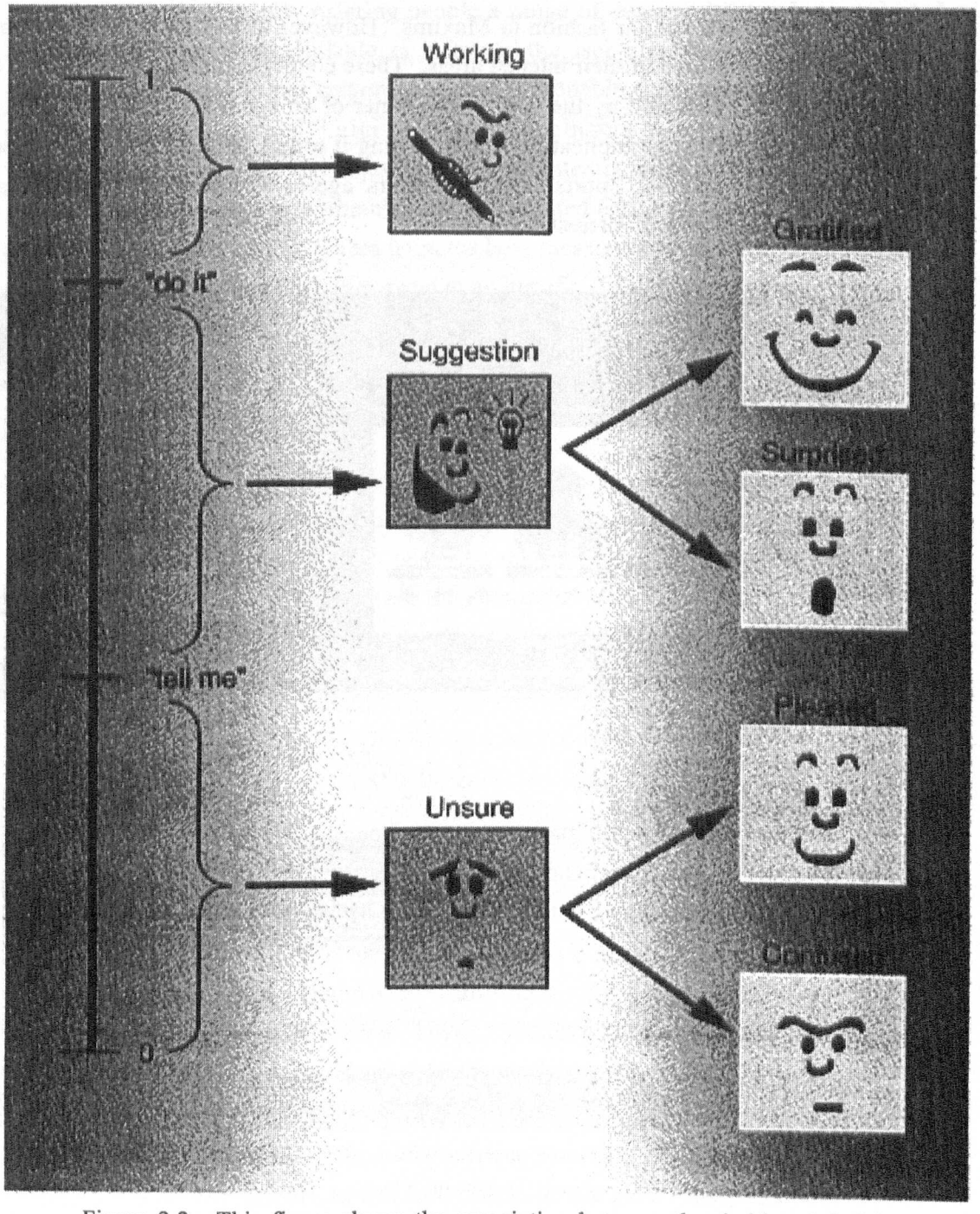


Figure 2.2. This figure shows the association between threshold and facial expression. Depending on the user approving or disapproving the suggested action the system responds again with a different expression. Taken from (Maes, 1994).

NewT worked in a similar fashion to Maxims. However in this case the user created a separate agent for each of their interest areas. These could then learn by observing, or by being specifically trained by the user. The owner of an agent could then share it with other users of NewT by duplicating it and passing it round. Figure 2.3 shows NewT at work with the 'business', 'sports', and 'computers' agents asleep and the 'politics' agent being trained on topics of interest.

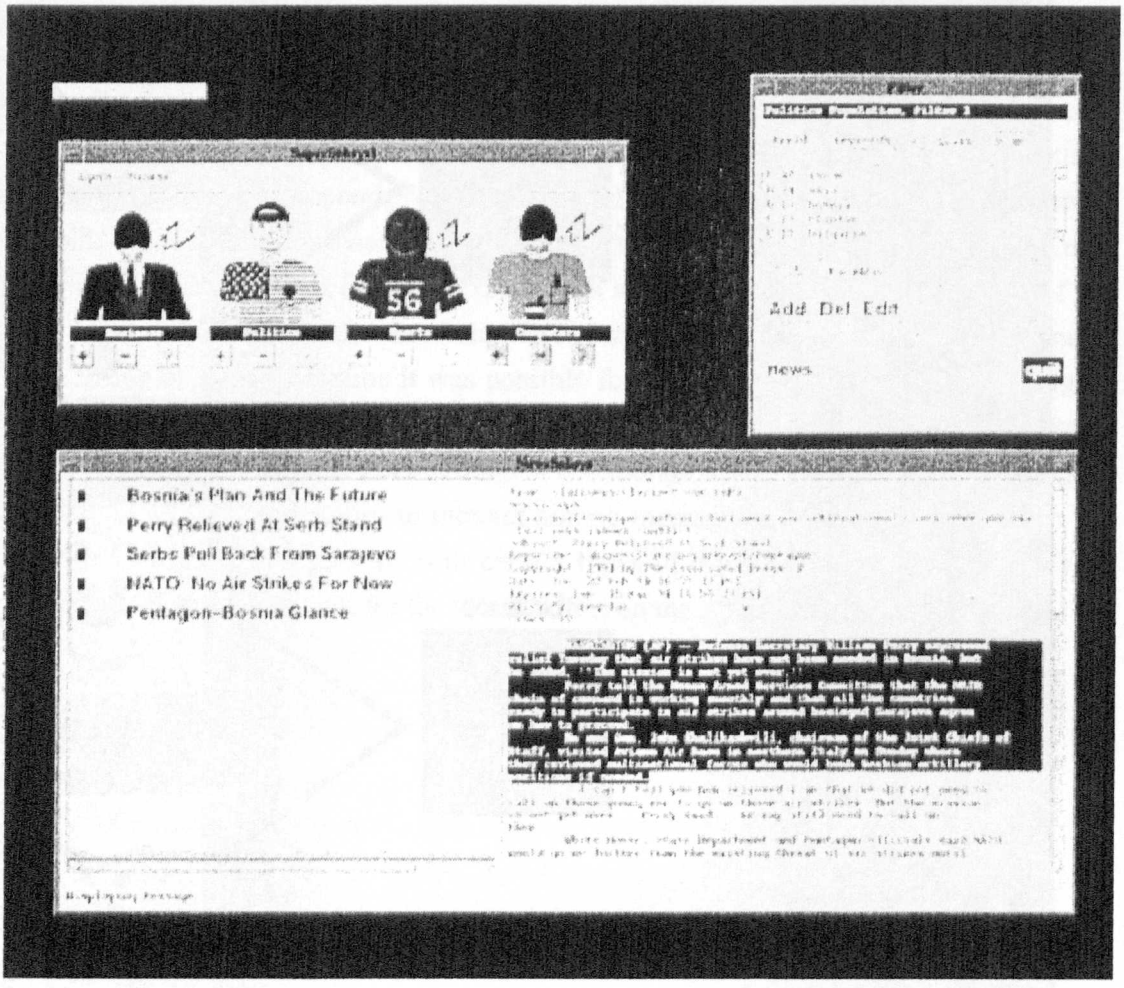


Figure 2.3 Showing the NewT interface with 4 agents. In this particular view the 'politics' agent is being trained. Taken from (Maes, 1994).

Finally, Luigi was an intelligent agent for scheduling meetings (Watt, 1996). The system attempted to address the problem of how to automatically manage the scheduling of meetings without requiring every person involved to be running a compatible diary system. Luigi enabled any user with access to e-mail to be able to arrange a meeting with

any other e-mail users. By offering people a range of dates and times Luigi selected a time when all users were available, or informed the user who has called the meeting of a conflict. Users who do not respond within a reasonable time get chased and when the meeting has been agreed Luigi will send reminders. Watt's paper talks about 'psychological agents' and discusses how humans interact with computers and 'intelligent agents' in a similar vein to Norman's work highlighted earlier. He builds on earlier work on groupware (Watt, 1993), where he notes how these systems fail when some users are required to do work for which they receive no benefit. The observations in this paper are similar those mentioned earlier about natural language and agent systems where Watt notes the "dissonance between these expectations and reality" in interacting with computers.

The systems we have looked at here share the common goal of assisting their users by automating mundane tasks. Each of the agents provides various levels of feedback to the user as to the current state of their task. As the user becomes more confident with the agent's ability to perform a given task the amount of supervision and monitoring can be reduced. Each of the studies mentions the need for users to be motivated in using the agents to be able to reap longer term gains.

*Agents that function autonomously*

When considering agents that function autonomously the systems which fall most easily into this category fulfil a role within a specific collaborative domain somewhere between being of general use to all users and being targeted at a specific user when the agent identifies a need. We will first look at two systems from the Advanced Technology Lab at Andersen Consulting, known as InfoFinder (Krulwich & Burkey, 1997) and ContactFinder (Krulwich & Burkey, 1996). We then go on to look at a brainstorming agent that provides documents from a database which are relevant to current discussions (Nishimoto, Sumi & Mase, 1996).

We start with the InfoFinder agent (Krulwich & Burkey, 1996) whose role was to hunt the web on behalf of its users. All interactions with the system were via e-mail in a very basic fashion. To start a search the user e-mailed the system a web page, in plain text format, related to their interest, using the subject field to give it a classification, i.e. 'agents'. InfoFinder would then add this to its current search of the web for all users. After a time (usually overnight) the user would receive an e-mail listing web pages found

matching the classification 'agents'. The user could then look at these pages and send them back to InfoFinder either as positive, '+agents', or negative '-agents', examples of that classification. The user could have as many different searches running concurrently as there were classifications. InfoFinder had one really awkward and one really useful feature. The awkward feature was that after the first 10 messages relating to any classification it would cease to learn (InfoFinder used an inducted decision tree) from all examples and instead only update the existing profile. This could never be reset. When I used it this meant that it became obsessed with Olympic Winter Sports, which I had no interest in. However the useful feature of the system was that once you stopped sending InfoFinder messages about a certain classification it would slowly cease to send you any messages about it. The system managed a collective web searching resource, sharing the results to cater for the needs of specific users. Although my own experience with InfoFinder was of some disappointment with its inflexibility it did, however, produce a result on a topic I was not able to find using any search engine.

ContactFinder is different in that rather than being directly approached by a specific user it monitors a group for a specific need it can identify and then attempts to satisfy it. ContactFinder works within a conferencing system (Lotus Notes) watching the messages posted. It analyses each message in turn to identify users seeking help, and the topic on which the help is needed. ContactFinder then searches its knowledge base for 'contacts' who might have knowledge in that area. The list of contacts generated is then sent to the originator of the request. In this way the system not only helps the originator of the request but might also relieve some of the burden on the rest of the users. However without the ContactFinder response being posted to the general conference this additional help is quite limited.

With both these systems there is a relevance problem. Users will be put off if the responses received from the systems are inaccurate or not useful. There has been some evaluation of both systems. In the case of InfoFinder there is a brief mention of an evaluation in one of the papers but no substantial detail. The evaluation of ContactFinder analysed 2893 messages, the system identified 611 questions for which it found 83 potential referrals (13.6%). Of these 83, 3 were discarded, 39 were approved by the contacts themselves and 11 refused, giving a claimed 78% success rate (Krulwich & Burkey, 1996). Anecdotal evidence I received from a user of ContactFinder was that the contacts suggested, although they would be classified as relevant, were often not very useful at all.

The "Outsider Agent" described in (Nishimoto et al., 1996) has the role of injecting a different viewpoint into a brainstorming session. In this way it behaves in an opportunistic manner, like ContactFinder, but for the benefit of all users. The theory behind this system is that an outsider's presence at a brainstorming session can provoke original and different ideas. The agent's creators have tried to construct an artificial outsider. The Outsider Agent attempts a shallow understanding of the participants' discussions and uses this to extract 'relevant' information from a text database. The hope is that the extracted texts will help stimulate different ideas and viewpoints. The authors of this paper have themselves, independently, used the term 'Virtual Participant' to describe the part of their system that joins in the conversations. Although the ideas behind this system seem pretty sound they have never conducted a serious evaluation of the system (so far they have only tried it with 4 users working on a staged problem). The main difficulty with such a system, and with the Virtual Participant in this thesis, is the problem of relevance. Users will rapidly be put off by a system that does not produce truly relevant responses - a problem common to all the systems discussed in this section.

*Collaborative filtering systems*

Collaborative filtering systems are not usually referred to as agent systems although they do fulfil some of the same roles. Users delegate their filtering task to a system which then functions autonomously, a cross between InfoFinder and Maxims. A discussion of these systems can be seen in (Resnick & Varian, 1997) where details of some of the important social problems characteristic of these systems can be found. Many of the systems have shortcomings affecting what the user is presented with and leading some users to receive an unbalanced benefit with respect to the effort they put in. This imbalance affects how much and how long users spend with the system. These problems need to be taken into consideration in any collaborative or group system and have been noted previously in (Watt, 1993) and (Grudin, 1994).

The Knowledge Sharing Environment (KSE) (Davies, Stewart & Weeks, 1998) is a good example of a recent collaborative filtering system. The intention of KSE is to provide a user group with an environment which enables them to more easily share both explicit and tacit knowledge. The KSE approach gives each user their own personal agent. Each agent manages a user profile intended to "model that user's information needs and interests". The KSE agent acts as an information management system with the users choosing to add documents which they find of interest. At the point of addition a user

can annotate the document and select which interest groups they think the information should be passed to. By making the users responsible for pre-filtering the information added to the database this guarantees high relevance of the information. The small groups of known users seems to have served to increase and maintain the usage levels.

Another relevant example is the recommender system 'GroupLens' (Konstan, Miller, Maltz, Herlocker, Gordon & Riedl, 1997). Recommender systems are a specific example of collaborative filtering systems, although the distinction is a fine, if not irrelevant, one. The most interesting point raised by Konstan *et al.* is the difficulty involved in getting users to give feedback, and what motivates them to do so. They describe a scenario where a small number of users are reading the majority of messages and providing the ratings. The rest of the users then benefit from this. Konstan *et al.* then go on to imagine a period when one of the users providing the majority of the ratings is away for a while. The user comes back to discover that articles are now already rated and their contributions are not really necessary - or that many fewer articles have received ratings as they were doing the majority of the work. Either way this dichotomy between those who do the work and those who reap the rewards causes the useful life of many collaborative systems to be short - often of the order of months. There is no clear cut solution to the problems these systems highlight, and solutions which work in one instance are not necessarily transferable to others. Being aware of these possible pitfalls helps to warn systems designers of what may await.

## 2.4 Tertiary Courseware

Mayes and Neilson (1995) define what they term primary, secondary, and tertiary courseware in terms of the tasks teaching systems are to play in what they call the "learning cycle", already shown in chapter 1, figure 1.1. Mayes and Neilson state that "most learning technology has been directed towards the presentation of content, rather than the answering of questions, or the opportunity for discussion and reflection". Because of this they believe that the majority of systems focus only on the content delivery mechanisms which make up "primary courseware".

Mayes and Neilson (1995) describe three systems which they classify as "tertiary courseware". The first is "Answer Garden" (Ackerman & Malone, 1990), the second is "Engines for Education" (Schank & Cleary, 1995), and the third is their own system "Answer Web" (Smeaton & Slater, 1994; Smeaton, 1995). The Answer Web system has

since been replaced by a more recent project called "The Vicarious Learner" (McKendree, 1996; Lee, Dineen & McKendree, 1997; McKendree & Mayes, 1997; McKendree, Stenning, Mayes, Lee & Cox, 1997). By studying other systems designed to support dialogue their strengths and weaknesses can provide pointers for the Virtual Participant.

The first, and oldest, tool is Answer Garden (Ackerman & Malone, 1990; Ackerman, 1994a) which Ackerman terms an 'Organisational Memory' and although not specifically intended for teaching, others have put its ideas to that use, i.e. 'Answer Web' (Smeaton, 1995). The metaphor behind the 'garden' is that of exploring a space in which different areas are on different topics, but this information is not fixed, it grows over time with new experiences and information. Answer Garden, along with its sister applications LiveDoc and Assist, were used as ways of managing and sharing information within groups of people. Answer Garden was tested in the domain of X Windows programming, with the simplifying assumption of distinct groups of experts and information seekers. The intention was for users to explore the garden, either simply browsing, or looking for answers to specific problems that they had. When they found the relevant topic they would peruse the information to see if a similar problem was already stored. If not, they could then e-mail the domain expert. The end result was that a new problem and its solution would then be added to the garden. The most immediate problem with this system was that users would identify the topic area in which they were interested and then just mail the domain expert directly rather than taking the time to search for an answer. The goal of reducing the load on the experts was scuppered by users looking for quick answers with no effort. In Answer Garden 2, (Ackerman & McDonald, 1996), Ackerman attempted to address this problem with the use of a 'escalation agent'. The agent was programmed to follow a number of rules about how to take a query forward. Eventually the user would get to the point of asking an expert, but they would have been taken through various other methods of solving their problems first.

Figure 2.4 is taken from (Ackerman, 1994b) and shows an example interaction with the Answer Garden system. The background shows the tree navigation structure with the foreground showing the user narrowing their search from 'The X Window System', to 'Using an X Application', then 'Finding an Application'. Finally, unable to find any answer to their question, the user e-mails the expert in that area.

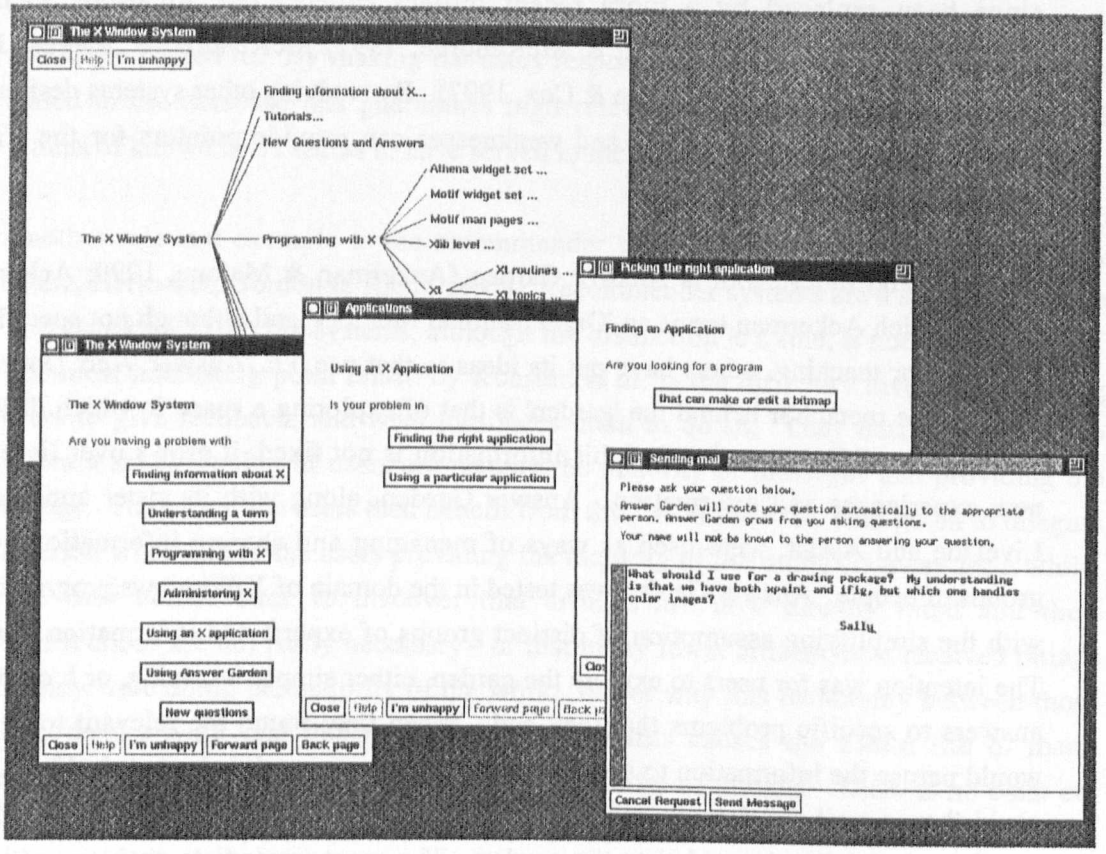


Figure 2.4. The interface to Answer Garden, from (Ackerman, 1994b).

In his thesis (Ackerman, 1994a) Ackerman analyses in detail the different types of Answer Garden user. From these studies Ackerman was able to categorise his users into three basic groups, "tire-kickers", "intermittent users", and "heavy users". His study highlighted some of the problems in getting people to use the system. To become a user of the system you must make two changes to the way you work. The first is task-related and involved the need to make use of additional software. The second was culture-related and required a change in your working practices. Either of these would be enough to put many users off.

In 'Engines for Education' (Schank & Cleary, 1995) Schank explains teaching theories, especially case-based teaching, and how a change is required in the way that schools teach children. The 'Engines for Education' are based on a technology known as 'ASK systems' (Ferguson, Bareiss, Birnbaum & Osgood, 1992), a name coined from their conception as systems to which you put questions and which provide answers in the form of a story, an Aesopic dialogue. The argument behind this approach being that students

learn best when that learning is contextualised. Critically, ASK systems are organised collections of stories, which Schank sees both as a fundamental form of knowledge, and as being ideal for teaching.

ASK systems can be thought of as a highly structured and interconnected database of stories and information. Almost all the actual systems are multimedia in nature, and while originally this really meant text and video clips more recent systems go beyond this with a role play, game-like, environment. Each system is tailor-made to suit one specific task and uses video clips of experts, whose time is often not cheap. However all the systems have a number of core elements in common, although these are modified for the chosen teaching approach and domain. Importantly, in each system the designers have accounted for the motivation levels of the intended users.

In the ASK systems designed for industry the intended users are generally highly self-motivated and so the learning tasks need not be disguised. In designing systems to teach children history the real task may be disguised inside a different task, for example that of producing a news story for TV. This is a kind of 'Trojan Horse' model of teaching, or as Schank terms it, "incidental learning". This allows the instructional designers to piggyback the motivation for the students to learn on to their motivation to make a TV news story. The shell system was designed to support this task by providing the students with all the background materials necessary to investigate the event about which they are intended to learn. This shell incorporated an expert critiquing element to guide the students in the form and content of their report, helping them to learn more about the subject area.

With the systems designed for industrial use the need for guiding tasks is eliminated and it is up to the student which different aspects they wish to learn and investigate. These systems are intended for students to use on an individual basis as they see fit. When they initially come to the system they start at the very top level of what is known as the 'zoomer'. This is a navigation tool which allows the students to focus down rapidly on the sections in the materials which are most relevant to them. In the system designed for children this process has been done for them.

The 'Engines for Education' CD ROM (Schank & Cleary, 1995) is an ASK system in its own right. When you first start the system you receive an introduction from Roger Schank and you are then left with a selection screen of viewpoints from which you can

choose how you view 'Engines for Education'. The four most common themes are presented. By selecting the zoomer from the navigation menu you are presented with the full ten choices, shown in figure 2.5 top left. If we select 'New Media' we are then presented with the choice of four stories the system would like to tell us, shown in figure 2.5 top right. At this stage you may also opt to view an outline of the contents tailored to this view point, shown in figure 2.5 bottom left. Once you have selected the first story you want to see you are taken into the ASK viewer, shown in figure 2.5 bottom right. It is also possible to jump directly to any story within the system through the navigate menu. There are approximately 100 stories within this system.

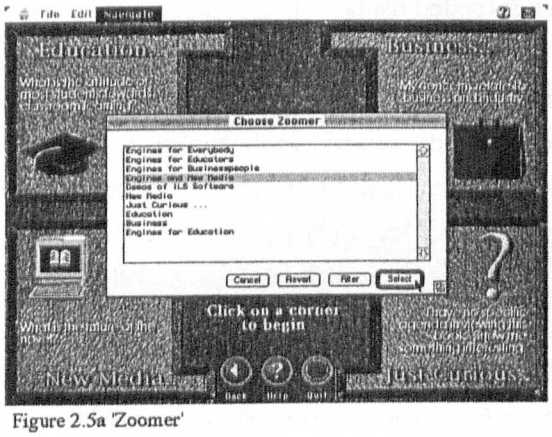


Figure 2.5a 'Zoomer'

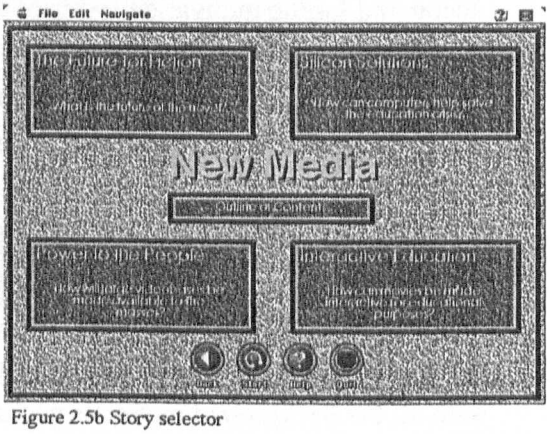


Figure 2.5b Story selector

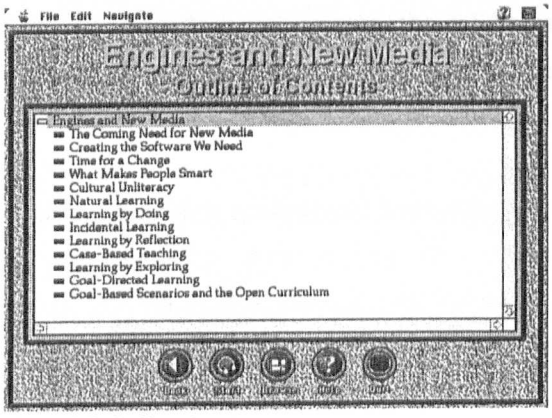


Figure 2.5c Outline of contents

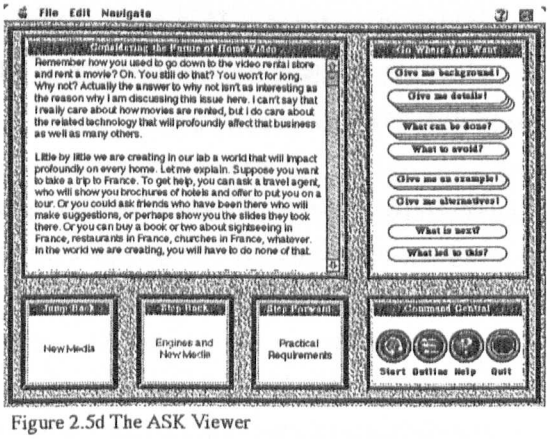


Figure 2.5d The ASK Viewer

Figure 2.5. Screen shots from the 'Engines for Education' CD ROM ASK system, (Schank & Cleary, 1995). a) The top level and zoomer, b) The theme "New Media" and associated stories, c) The outline of contents from the "New Media" viewpoint, d) The ASK viewer.

Once inside the ASK browser the student has a number of choices available to them, shown in figure 2.6. For navigation purposes they can choose from the three buttons in the bottom left allowing them to 'Jump Back' to the highest level which is 'New Media', 'Step Back' to the previous option which was 'Engines and New Media', or 'Step Forward' to the next story which is 'Practical Requirements'. It is possible to view the core of the content by continually pressing 'Step Forward'. Or the student can choose to navigate by the 'Go where you want' sidebar asking questions like 'Give me background' or 'What can be done?'. The height of each button indicates the number of options behind it. This allows the user to move through the system following their own approach and in their own way. One immediate downside of this particular ASK system is the lack of a bookmark facility to enable you to come back to the same place with the same options.

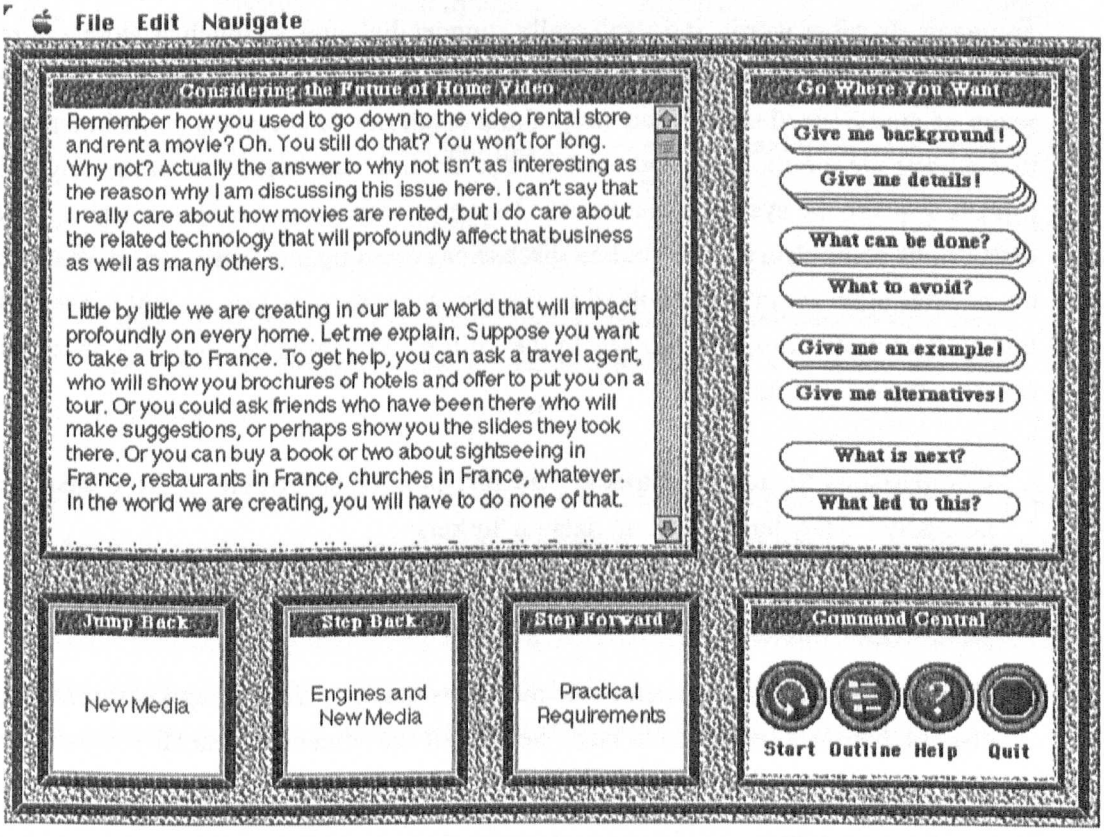


Figure 2.6. The ASK viewer from 'Engines for Education' (Schank & Cleary, 1995).

The final system to look at is "The Vicarious Learner" (McKendree, 1996; Lee *et al.*, 1997; McKendree & Mayes, 1997; McKendree *et al.*, 1997). The task of the Vicarious

Learner project was to address the problem they saw with the majority of electronic, specifically web-based, systems providing what they termed *primary courseware*. The Vicarious Learner, like all systems discussed in this section, was intended to support discussion amongst students as *tertiary courseware*. The theory they were expounding was that observing other learners and their discussions about certain problems may help the student learn. This goal was realised by storing student discussions and linking them into on-line course materials. The project overview can be found in (McKendree, 1996), and details of the initial studies in (McKendree & Mayes, 1997; McKendree *et al.*, 1997). The main problem with this approach is that while the idea is sound, supporting dialogue, the practical implementation is wrong. The students are in no way motivated to start or continue discussions and instead the system simply links stored discussions to existing material extending a *primary courseware* approach rather than a *tertiary courseware* approach. In other words, it doesn't really *support* dialogue, so much as store and present it. A good review of the project can be found in (Lee *et al.*, 1997) which compares two years of study. In the first year only about a quarter of the class contributed, but the discussions were quite deep and varied. The second year improved on the original HCI aspects to make the system easier to use. However they made the fatal mistake of making discussions compulsory. The student discussions dried up and were very shallow. There are several other variables involved in the two separate user groups, but it seems likely that student motivation was the key to the failure. The project appears to have failed to consider the following:

- Comparisons to normal classes: Do one quarter contribute and the rest observe normally? They have given no data on 'lurkers'.
- No consideration of student motivation: Making discussion compulsory does not serve to motivate anyone.
- These students are not distance students: It is quicker, simpler, and easier for them to discuss their problems face to face. Several of the student comments support this.

Rather than consider any of these factors they have chosen to use 'task-directed discussion' methods, (Lee *et al.*, 1997), to try and promote more discussion earlier. I believe that without the consideration of the above points this is doomed to failure. I will come back to my experiences with these problems later in this thesis, in chapter eight.

## 2.5 A summary of case presentation systems

This chapter has presented various systems with a range of applications all of which have lessons that can be learned from. This final section reviews the four principal methods for the presentation of cases; a summary is given in table 2.1. The approaches could use the same set of cases, but each has its own presentation style. This taxonomy appeared previously in (Masterton, 1997).

***Taxonomy of case presentation systems***

| System | Interactions with users | Format of cases | Retrieval | Maintenance and maintainer |
|---|---|---|---|---|
| FAQ | Passive | List of cases | Manual | Low - single human |
| Contact Finder | Active | Hidden database | Automatic | Low - automatic |
| ASK | Users choice | Highly structured and interlinked | Manual exploration and automatic guidance | High - single or multiple authors |
| Vicarious Learner | Passive | Linked to relevant course materials | Manual | Medium - tutors or students |

Table 2.1. Summary of case presentation systems, how they interact with users, the format of the cases, how the cases are retrieved, how much and who maintains them.

First, there is the idea of a 'frequently asked questions' list (FAQs) (Hambridge, 1995), a case-base of questions and answers stored as a list. To make use of this passive list the user must comb the list manually, or use a search tool, to find something matching their problem. The list is maintained by a human who collects and saves these examples from an associated electronic conference. Using FAQs is a common approach in a number of areas, however there are a number of shortcomings for the user of lists of FAQs:

- Can not find it (the FAQ itself, or a relevant question).
- Previously read the FAQ, but can't remember the contents.
- Found the information after I had already solved the problem.
- It is often easier and quicker to ask.

Much to the frustration of the maintainers of FAQ lists the fourth point commonly arises in so called 'newbie' questions to discussion forums. This same problem was reported by Ackerman (Ackerman, 1994a) who found some of his users held the same philosophy. Open University course tutors who have included FAQs in course materials or archive important past discussions commonly complain that 'the students never read them'. What we observe here is the principle that users are often unwilling to exert excess effort (Watt, 1993) and will take the easiest route, i.e. asking the expert.

Secondly; 'Contact Finder' (Krulwich & Burkey, 1996), actively watches for questions in a conference, matching them to its hidden case-base of previous examples. When a match is found it uses this to direct the questioner to someone who may be able to answer their problem. Once set up the system is automatic, requiring no human intervention. This is a kind of 'Recommender System' (Resnick & Varian, 1997) similar to many which have been developed for collaborative filtering, and which often have the following social implications:

- **Incentive problems:** It is hard to gather recommendations, and guarantee the quality of those recommendations.
- **Personal privacy:** The more information about a user the better the recommendations, but users are unlikely to want these profiles widely shared.

Thirdly, and specially intended for teaching, are 'ASK systems' (Ferguson *et al.*, 1992). When using an ASK system students are able to switch between active exploration of the case-base, to being passively guided through in the way previously set by the tutors. ASK systems require a large amount of maintenance from the tutors, especially in the setting up stage when a highly linked case-base is created with guided steps for the students.

Finally there is the 'Vicarious Learner' (McKendree & Mayes, 1997; McKendree et al., 1997). In this system students work through hypertext course materials and discuss them on-line. Relevant discussions from previous years are linked into the course materials by the tutors. The students then have access to these discussions as extra materials.

The Virtual Participant draws from the pedagogical aspects of the 'Vicarious Learner' and the automated aspects of 'Contact Finder' to create a system which provides FAQs with some of the structure found in ASK systems. Using the same categories as shown in table 2.1 the first Virtual Participant prototype was intended to provide active user

interaction, a structured hidden database, automatic retrieval on matching, but manual retrieval from then on, with tutor maintenance. This prototype system was then evaluated and the results of this evaluation fed back into the development process of the project. The final system had a public database, manual and automatic retrieval of everything, and the maintenance burden was more evenly shared by encouraging student input.

## 2.6 Summary

The work which has been reviewed here has been chosen purely on the basis of its relevance to, and influence on, the development of the Virtual Participant concept. The literature that has been reviewed highlights many facets that must be considered and emphasises the fact that with all these approaches it is not the technical problems that are the hardest, but, rather, those involving people.

The chapter started off by looking at agents in teaching and Chan's taxonomy of them. We have seen how the Virtual Participant concept extends and transcends the categories he has suggested. Moving on from educational agents we studied how users might interact with the agents themselves and some of the pitfalls that await agent designers. It is never possible to satisfy everyone, so agent designers have a difficult task of trying to cater for a large number of tastes. Successful systems show themselves to be flexible and provide 'out of the box functionality', so to speak, in the rapid way with which they establish their usefulness.

The ideas behind tertiary courseware propose that dialogue amongst students is important to learning. Supporting this process is of the utmost importance and it is clear that the current systems have not quite cracked it. We must learn from these approaches, considering what motivates students to discuss issues in the first place. Understanding this process will help to produce better applications.

Finally we have summarised the properties of four different systems specifically designed for case presentation. These properties provide a starting point for selecting various development options for the Virtual Participant concept. At the end of the first study the properties were again used to assess success in each area and to provide points where the second prototype needed to be developed. The evaluation of the first prototype can be found in chapter 5, with the specific property breakdown in section 5.6.

Comparing the case presentation approaches of the Vicarious Learner, ASK systems, ContactFinder, and Answer Garden after tutorial materials have been added, there is one clear key difference, the use of context. Context is a major problem and something the FAQ approach is the worst at, hence the specific problems it faces. All the other systems contextualise their information:

- The Vicarious Learner attaches discussions to relevant parts of course materials.
- ASK systems have been manually structured to provide related information to pre-stored stories.
- ContactFinder provides information in response to queries.
- Answer Garden links relevant garden sections to tutorial materials.

We have reached two conclusions, context is vital, and the perceived benefit must outweigh the perceived effort. Contextualisation is behind Schank's theories on case-based teaching and learning (Schank & Cleary, 1995). Students can learn from observing real-world case studies about certain phenomena, rather than just learning the theory behind them. Learning from cases is a common theme in management training (and it is central to Law where certain cases set precedents that can then be used to argue for or against other similar cases). It is also the idea behind the Vicarious Learner (McKendree et al., 1997) where students can learn by vicariously observing the dialogues of others - behaving rather like 'lurkers' in conferencing systems.

The aim of this study is to produce a practical system so that the Virtual Participant concept can be tested. In doing this the idea of taking cases and setting them in context can have its influence and applicability properly assessed. The remaining chapters take you though the development and assessment of the initial prototype, the second study, and on to the final conclusions.

# Chapter 3 - The First Prototype

## 3.0 Background

In the previous chapter we looked first at agent technologies as an approach to interacting with students, and then at tertiary courseware to deliver information useful to the students. The idea behind tertiary courseware is to support discourse as part of Mayes' learning cycle (Mayes & Neilson, 1995). The research discussed in the rest of this thesis came about through working with the Open University MBA elective course B882, 'Creative Management' (Martin & Henry, 1993). This course was the first in the MBA (and one of the first in the University) to adopt text-based electronic conferencing as part of its teaching approach. The course provided an unrivalled resource of previous dialogue as all public student discussions from 1991 onwards had been stored, forming an ideal starting point and seed for a tertiary courseware system.

The initial idea of the Virtual Participant was met with interest by the academics responsible for the course, the 'course team'. But to build a good foundation for the study it was necessary to look in more detail at support issues for students and tutors in electronic conferencing. Immediate observations and discussions with tutors and course team indicated that the most significant problems were those of:

- Tutor overload: Too much to do with too little time.
- Need to get all students to participate: To increase overall benefit and learning.
- Poor quality of content: Off topic discussions, students not knowing what to ask.
- Same questions asked each year: Similar materials provoke similar questions.

These points can be seen as the high level goals that we were looking to address. The tutors and course team did not have a clear idea of how these goals could be addressed and what kind of approach would prove successful. However there was a general agreement that something, *anything*, that could help or assist would be a good thing, with the proviso that it did not negatively impact on any student. The initial ideas were inevitably somewhat idealistic, mostly because the first prototype was an exploration vehicle to discover what could be done, and what students and tutors thought of the general approach of tertiary courseware. By reviewing the successes and failures of the initial prototype the ideas could then be refined or even redefined in the form of a second prototype. The details of this second prototype can be found in chapters six and seven.

From the starting point of a large amount of raw discussions from past years it was necessary to process them in some way that they could be both usable and retrievable. Secondly it was necessary to consider some type of delivery mechanism for the processed discussions, a mechanism which would interact with the students and promote dialogue. Because of this split in tasks this chapter naturally falls into two parts. Part A looks at our existing resource of past discussions and whether they could be used to generate clusters of similar discussions and automatically identify frequently asked questions. Part B then looks at delivery, how the results in part A could be combined with the agent technology reviewed in chapter two to put together a working prototype. Figure 3.1 shows the dependencies within this chapter indicating how the approach taken in part B is dependent on our choice of a tertiary courseware approach and the results of the clustering of past discussions from part A. The details of the implementation depended on the ability to successfully index and retrieve data.

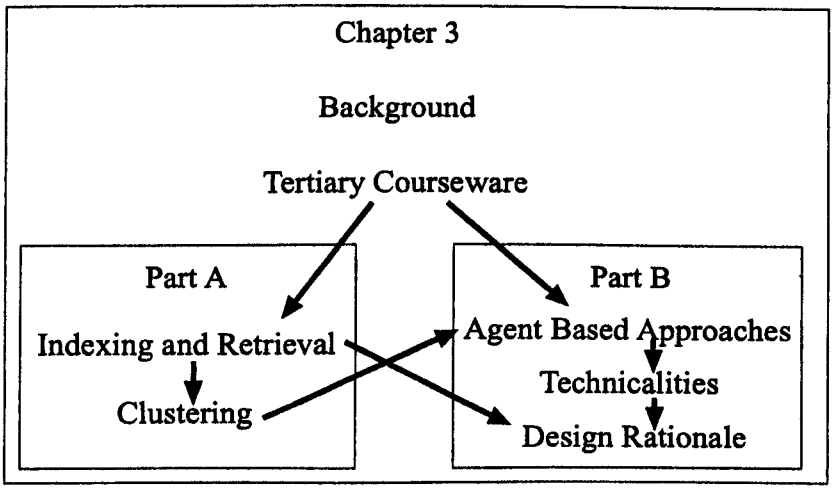


Figure 3.1 Diagrammatic guide to the dependencies in chapter 3.

## Part A: Past Discussions as Frequently Asked Questions

In this part of the chapter we look at the archived past discussions for this course, and how, from these, we can identify frequently asked questions.

The use of FAQs requires certain assumptions. Although these are particularly relevant to Open University courses we feel that they are generally applicable to all courses.

- Premise 1) The course materials change only slightly, if at all, each year.
- Premise 2) The assignments cover similar issues each year.

Given these two premises.

- Assumption 1) The discussions each year will be on similar topics.
- Assumption 2) The words and phrases used to discuss these topics, and the names of the topics themselves, are contextually dependent on the course materials, and therefore will be identical to those used in the course materials, so the same words will tend to occur each year.

Assumption number 2 is very important and the basis of much of this work. If the discussions have similar context each year, and we can identify that context, then we can match current discussions with past ones based on this context. Schank's theories (Schank & Cleary, 1995), as but one example, imply that contextualising the information to the students' current needs would give an improved learning experience.

In the next two sections we look in detail at the natural language side of our problem. Firstly we look at the existing approaches to indexing and retrieval. There are two approaches here, the statistical approach of natural language processing, and the hand-crafted approach of case-based language understanding. Then we go on to look in detail at the archive of discussions and ways of turning them into a base of frequently asked questions.

## 3.1 Indexing and retrieval

In this section we are looking into the natural language problem of identifying current topics of conversations so that we can link them to past topics of conversation in a conferencing environment used for teaching. A principle of both statistical natural language processing and case-based reasoning work is that indexing and retrieval go hand in hand. We cannot consider the identification and retrieval of topics without considering how the topics are first indexed.

The most naive and simple approach would be to take every message from every previous year, store them all in a database and allow free-text searching. There are many ways of indexing free text documents (Salton & McGill, 1983) and one could be chosen. Unfortunately this completely ignores the context of these messages. The majority of messages are incomplete without the context of the surrounding thread of discussion. Alternatively, it would be possible to store whole threads of discussion and retrieve them in the same way as individual messages. This presents a classic case-based reasoning problem. If you have a number of very similar cases (which we are assuming we will have) then do you keep them separate or combine them into a meta-

case? Case-based systems are often hand-crafted for the topic area, and this decision is never a clear one (Kolodner, 1993). Taking a step back for a moment, perhaps we are looking at this problem the wrong way. What do we really need in our database? To answer this question properly it is necessary to look back at the course itself.

Course materials cover a number of topic areas, and although they may be interlinked the students do come across them as separate topics and at separate times. Perhaps this is the key to solving the problem. Indexing messages or threads is not the solution, nor even the problem. The problem is to identify common *topics* across the years and find a way of both indexing and retrieving them.

In the next section, I'll describe the exploratory work we conducted on our database of previous years' discussions. First I'll look at two methods of indexing which could be applied to our database.

Put simply, there are two main methods of indexing which could be appropriate: fixed and dynamic (these are sometimes referred to as absolute and relative). Fixed indexing requires us to select the indices we are going to use initially and stick with them. This approach requires the initial overhead of processing each of the indexed items and selecting terms, single or co-occurring words, for it. However this process can be done by hand and a good indexer will be able to choose index terms not included in the document but which are relevant to it. When the system is running it has minimal overhead in terms of retrieval. For dynamic indexing things are not quite as straightforward. There is no initial indexing of the cases, and all the work is done during the functioning of the system. Rather than fixing the indices by which the cases can be retrieved instead you use all the terms available. By using more sophisticated technologies in the form of electronic dictionaries and thesauri, you may even be able to find related terms. In this way fixed indexing is dependent on the terms chosen coming up again. Dynamic indexing, on the other hand, just requires a certain number of terms to be found before the document is retrieved. In both cases you would remove a 'stop-list' of words (like 'the', 'and', 'of', 'it', 'he', 'she') which add nothing to the context.

The downfall of dynamic indexing is that there is no way of guaranteeing that the most important terms occur and you run the risk of retrieving documents which are not relevant but contain a number of similar basic words (i.e. words which are irrelevant to the context of the document, but not on a 'stop-list'). The downfall of fixed indexing is that vital terms may have been left out or incorrect ones chosen. The problem with all statistical methods and those based on the occurrence or non-occurrence of terms is that they will be able to tell you that two articles are about the same thing, but not that one is

for it and the other against it. But this is not a thesis on natural language processing and for the purposes of this project any approach that provides a reasonable solution is a good approach - and the simpler the better. However it is important to be aware of the existing technologies their advantages and limitations. I chose to go ahead with fixed indexing as this is the fastest and simplest method, which avoids any programming, and puts me in control of the indices. I argue that the Virtual Participant should work irrespective of indexing method, a better method would just match more effectively, but would not make a fundamental difference to the success of the system. This argument is borne out by the conclusion to the studies presented in chapter eight.

Finally, before moving on to the details of creating the case base, there is one other matter to deal with. A large amount of the content of a thread of discussion is simply agreement and re-statement - this serves as an additional point against using unedited messages or threads. If we are going to look at producing and indexing topics of discussion then to maximise the content it is going to be necessary to create summaries. By creating summaries of the key points of a discussion, or at least the key messages, the users will be getting high quality feedback from the system. As I have opted for a fixed indexing approach I will be generating the indices either manually or in an computer-assisted way. This means that these summaries are important only for the presentation to the learner. If another indexing approach was chosen then that information which is cut from the summaries would still be very important to indexing.

## 3.2 Identifying common cross-year topics of discussion

One of the central assumptions of the first prototype is that of repetition. Each year the course materials are the same, the assignments are the same, and therefore the discussions and the problems tend to be broadly similar. When I first met the tutors who work specifically with the students on-line I was told 'no - every year is different'. I started out with six years worth of electronic conferencing messages for this course and the next few sections discuss the work I did analysing them. I have also read just about every message in those archives and it does seem apparent that the same issues arise time and time again. The concordances also seem to back up this view.

The 'Creative Management' course consists of four blocks, each eight weeks long. At the end of each block the student is required to submit a tutor marked assignment (TMA) which counts for 12.5% of the course, an exam making up the final 50%. The course is designed to take the students through various levels of the creative process. The assignments are based around students' own experiences and their current job, and this does make for some unique problems every year. It also means that by block 3

there is a rapid divergence on the topics discussed over previous years. In block 1, 'Perspective', the students investigate what is creativity, and what kind of creative person they are. This involves taking two personality tests which always provoke a lot of discussion. In block 2, 'Techniques', they are introduced to over 200 techniques for creative problem solving, which provokes a lot of discussion on what the techniques are like and how useful they prove. They are not expected to learn all these techniques, but rather they provide a resource from which the students can choose approaches which best suit their current problem. By block 3 they are applying these principles to their own problems and discussing them, and block 4 is a personal project. So once they reach block 3 the discussions have usually diverged so far that there really is little in common between the years. Because of this the work and investigation has concentrated on the first two blocks, and that is where we start looking at common discussions across years.

### 3.2.1 Making the threads

Initially I started off with the raw archives of every year's conferencing from 1991 to 1996. During these years B882 conferencing was conducted via a proprietary Conferencing System (CoSy) provided by the university's Academic Computing Service (ACS). CoSy ran via a vt220 (teletype) interface and was very similar to the original Usenet news readers. Its success amongst OU students rested on an off-line reader for Windows 3.1 called WigWam. WigWam stored each subscribed conference as one text file, which could then be used as the source for my archives. The first task was to take these raw archives and extract the different threads of discussion from them.

The CoSy archives provided enough information to link and extract the threads of discussion. Most of this came from the details of which message number was a reply to which other message number. However not all students managed to reply in this way and would sometimes create new threads by mistake. These isolated replies were matched up with their threads via the subject and date fields. This produced the initial set of threads for each year. Also because the original moderators had chosen to split the discussions on different TMA's between different conferences, the threads in these conferences were collected together. In practice, I used only the discussions from the three years 1994-96, as before then the amount of discussion was very limited.

If we look in more detail at the actual threads there can be wild variation in the number of messages in each thread. For example, TMA 2 in 1994 was pretty typical; it had 39 threads ranging in length from 1 to 24 messages each and a total of 209 messages. The

first quartile of these threads (10) only contain one message, the next quartile (9) contain only 2 or 3 messages. There are really only a few threads containing extensive discussion (for the detailed values of messages per thread for this and all years see appendix 1). To understand better what kind of discussions are taking place it is necessary to examine the contents of these threads. Looking at the threads containing only one message, these are mostly administration messages or students asking questions which are never followed up. Those which are two or three messages long can often be threads which didn't take off, could be dealt with easily, or which generated no interest amongst the students. The longer threads become more interesting but are not all necessarily about the course itself, but perhaps about missing materials, for example. The longest threads often contain a lot of 'me too', 'I agree', type messages through which the content gets diluted. Figure 3.2 takes the longest 16 threads from each year and looks at the length of each thread as a proportion of the total number of messages sent.

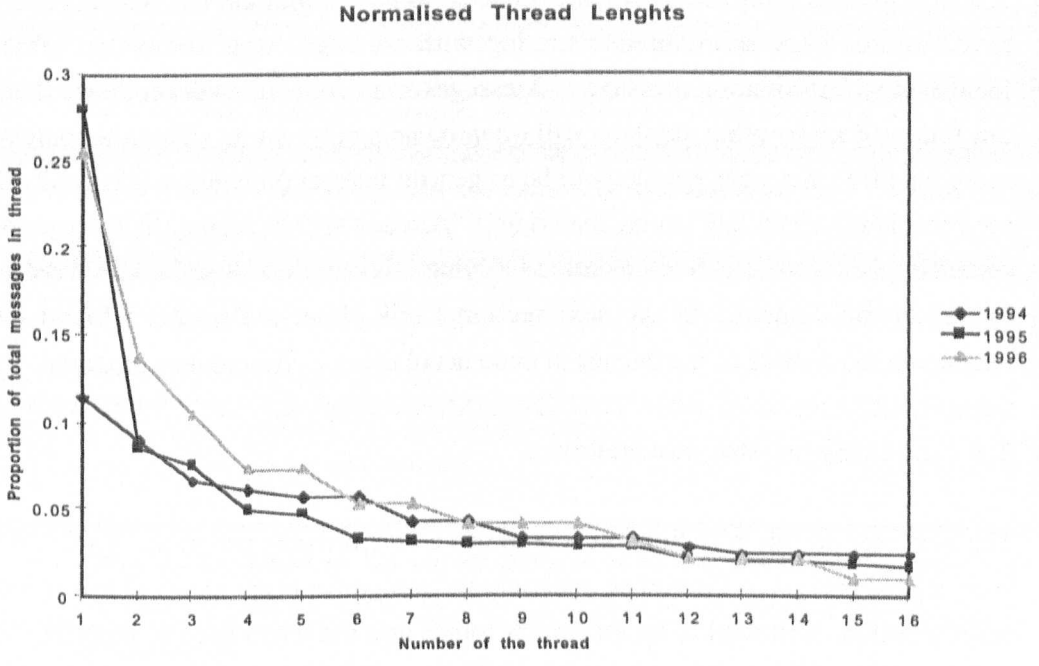


Figure 3.2. Lengths of threads in each year as a proportion of the total messages sent, first 16 threads displayed only.

This demonstrates that this pattern of discussions is broadly similar each year, even if the topics discussed are quite different. From the figures in appendix 1 you can also see that the median discussion length in each year was very similar, supporting this view. We can see that there are a few topics which the students really get into, with the

others dropping off proportionately. In every case about half the threads have 5 or more messages.

It is noticeable that in each year there are a small number of threads that have caught the interest of many students. These threads easily gain up to 20 messages or more. Apart from these, there are the announcement threads, generally one or two messages, which are just tutors giving out instruction details. Threads which consist of one message from a student seeking help are pretty rare, generally the tutors are very effective at identifying any student who has not been helped and trying to help them. This does lead to a few short threads where the conversation has not really taken off but at least there has been an attempt to address the student's concern.

*Summary*

A lot can be learned just from observing the size distribution of the discussion threads. The size gives a immediate heuristic guide as to how important the students found a given topic. Very short threads starting with messages from tutors can often be identified as 'information messages'. Messages from students with problems that are not followed up by other students will often be helped by tutors. Threads containing more than five messages are likely to be of general interest (although occasionally they are complaints about late course materials!). Already we are seeing the emergence of some simple heuristic guidelines for identifying relevant threads before we have even looked at the content. In the next section I talk about the results I found when examining the content of the threads in more detail using a concordance program.

### 3.2.2 Looking at the concordance

The next step in the investigation was to use a concordance program. So far, it has been argued that given the course materials and assignments are identical across years, the discussions will tend to be on similar topics and the terms used to describe those topics will be similar. By using a concordance program as an exploratory tool it is possible to look in more detail at different combinations of texts to find frequently occurring words or terms.

Initially I used a concordance program to look at complete years, and then at the threads within those years. The program allowed for the text to be both filtered and unfiltered through a 'stop list' of common words to be ignored. This helped focus on common terms to the discussions, rather than just common terms in the language. I used this same 'stop-list' in both the prototypes and it can be found in appendix 2.

Unsurprisingly the concordance shows that there were many common terms used that occurred in each year and that the same terms occurred cross years. These were terms like 'Creative Management' (the name of the course), MBTI (Myers Briggs Type Indicator) & KAI (Kirton Adaptor Innovator) (the two psychometric tests that the students take), and other terms which were variants on 'tma', 'tma 1', 'getting started' and so on. In figures 3.3 and 3.4 you can see examples of the words TMA and MBTI. By selecting the word in the index in the bottom window, the system shows the context in the concordance in the middle window, and the actual text it is taken from in the top window.

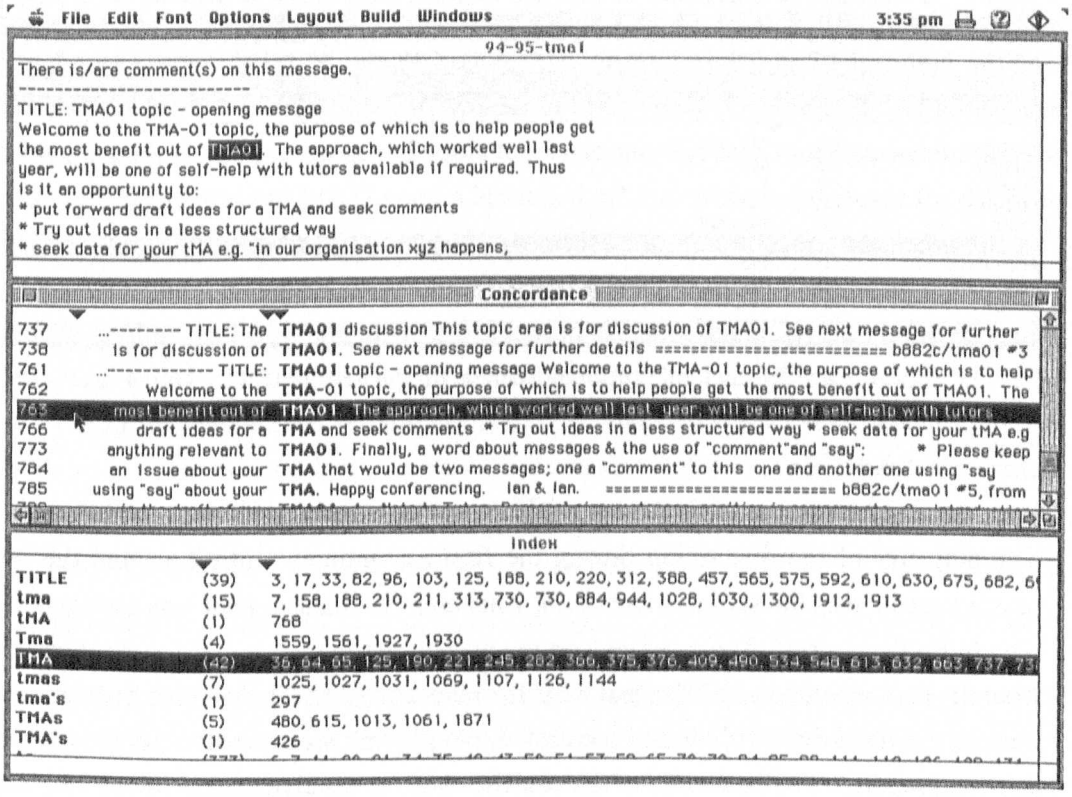


Figure 3.3: Extract of the concordance of the term TMA. Using the concordance program 'Conc', by John Thompson of the Summer Institute of Linguistics.

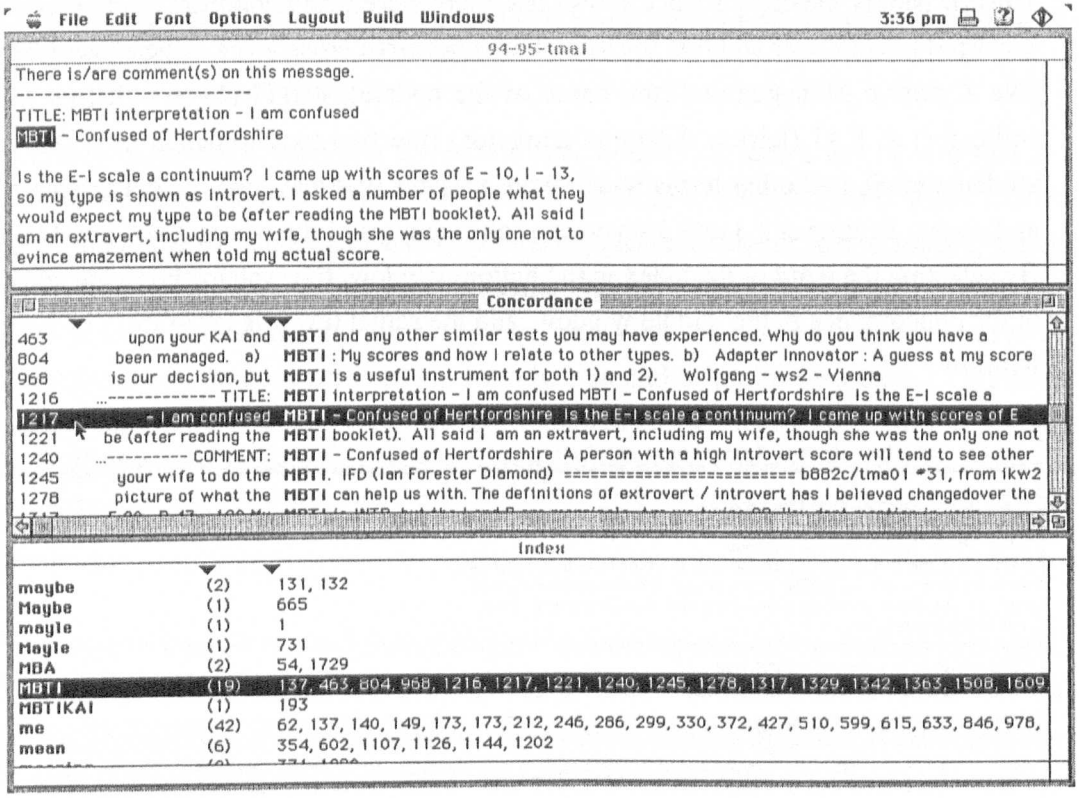


Figure 3.4: Extract of the concordance of the term MBTI. Using the concordance program 'Conc', by John Thompson of the Summer Institute of Linguistics.

The majority of the discussion during the first assignment centred around the two psychometric tests the students had to do, what creative management was anyway, and how to do the course. The questions on the psychometric tests mostly centred around what the tests actually tested for, and what the results meant? The creative management threads revolved around how to be creative, people that were creative, organisations that were creative and how none of the students were in creative organisations and how much they would like to move job![1]

The most interesting point that emerged from the concordance program was not that the same terms are repeated every year (this was predicted), but rather that threads could

---

[1]There is always a clear subset of students who are doing the MBA program to get a better job and they see B882 as the course to help them solve this problem. Another small subset are those students who are consultants or otherwise self-employed. The course does revolve around applying the techniques to your own companies and colleges that you work with and this proves difficult for some students.

easily be partitioned on only a few identifying keywords. The majority of topics that occurred could easily be categorised using a few terms that clearly identified it. At this point I went back to the course materials and looked at the contents page. The section and sub-section titles of each block strongly reflected the terms that occurred. This did not define every thread, for example 'how to do your TMA' and the 'psychometric tests' were rather different, but it did provide a number of seeding terms that could searched for with the concordance. These helped to identify each thread in great detail, and showed the contextualising influence of the course as a whole.

From looking in detail at the discussions it is clear that not all topics occur each year, and that the style of the tutors can have an impact on what is discussed and how conference use takes off. For instance, the problem of a self-employed student attempting B882 was only raised in two out of three years. Other effects are noticeable as well, such as the occurrence of some topics as sub-topics in one discussion. In one case a discussion about MBTI results branched off into what these meant for the self-employed and the types of creativity relevant to them.

Conferencing on B882 has always been optional and has often had relatively low uptake in terms of the students' participation. The years I have been looking at, 1994 to 1996, produced around 100 messages in the first assignment discussion conference each year. This is not a large amount of discussion but it is probable that the trends observed are applicable to many, if not all, other courses that use conferencing.

*Summary*

The use of concordances very quickly helps to highlight the common terms used to define different topics. A 'stop list' of common words also helps bring out more clearly those individual words that define topics. However the full meaning can sometimes be lost if simple words like 'not' are ignored. One common problem with AI systems is that they are often hand-crafted and this makes for great problems for mass take-up. However, identifying common terms that distinguish between threads on different topics indicates that a clustering program should be able to group the threads together, helping to find common patterns between years. The next section looks at clustering in more detail and assesses possible approaches. The concordance data provides useful indexing keys when it comes to the final creation of the database and we will therefore be returning to this later.

### 3.2.3 Clustering

Having decided to go ahead with looking for clusters of common topics across the discussions from previous years it was necessary to develop an approach to clustering. The previous discussions were already grouped by TMA and by year providing a file per year and per TMA. Then a program was developed to split these files up into their component threads. To see how easy it would be and to help me get a better grip on the contents I first attempted to cluster the threads manually. This turned out to require a lot of effort, reading most of every thread and attempting to classify it, then trying to link threads between different years. The task was laborious but in looking at the discussions in such fine detail it was easy to spot common terms used to describe similar problems across the years, again lending credence to the results of the concordance programs and indicating automated clustering as a viable way forward.

When looking at automated clustering we must consider two alternative approaches, symbolic and statistical. For the symbolic approach there are the induction algorithms developed by Quillian, either using complete data sets, ID3, or a cumulative approach, COBWEB (see: Luger & Stubblefield, 1997). These approaches produce decision trees that represent the distinction between different threads. Each thread would form an individual leaf of the tree. The closer together the leaves, the more difficult it is to distinguish between them and therefore the more similar they are. The decision trees generated by these approaches would provide sets of keywords which could be used to distinguish topics, and hence could be used to identify similar topics in new discussions. These sets would be made up from each word from each branch starting at the root node and ending at the leaf. Equally the tree could be 'pruned' at one point to indicate a topic. All the words in the branches from that point to the leaves could be used to describe that topic.

I decided not to go with the symbolic approach and to look more at statistical methods because they offer more explicit control over the effect of frequency and position of words if I chose. I was also looking for a technique which would enable me to make some kind of 'distance' measure between threads which could then also be used in the Virtual Participant system for retrieval based on how close new threads were to old ones. The decision tree approach is a more formal yes-it-is / no-it-isn't decision, compared to a distance function where a 'threshold' approach can be taken to matching.

There are various statistical taxonomic distance measure methods which can be used: angular coefficients, distance coefficients, Euclidean distances, and square block distances to name but a few. A good overview of these techniques can be found in

(Sokal & Sneath, 1963; Boyce, 1969). However, there is always an element of noise in the contents of threads and it will be necessary to edit the output of any algorithm. Even so, any kind of reasonable output would be a helpful start.

*Developing the clustering method*

The clustering approach that was used can be described in the following way. The approach can be compared to set theory, where set X is a subset of set Y if all elements of set X are contained in set Y. Two documents are said to be "neighbours" if the number of elements of one contained in the other is greater than a certain threshold. Note this measurement is neither normalised or reflexive. By setting the threshold to its maximum value all the documents will be in a set containing only themselves. As you lower the threshold value the documents will cluster to form bigger sets, until eventually you have only one set containing all the documents. The art in this method is knowing when you have approximately the right number of clusters of the right sizes. If you initially eliminate common words, using a stop list, from the documents you are comparing then as long as your corpus documents are related in some way then you are highly likely to get good results. When comparing documents from different corpora a problem will occur, as with all statistical methods, when terms common to each corpus are used to describe different things. However, we are dealing with a corpus where the documents are all related as having at least something to do with the MBA course on creative management, so this problem is far less significant. In this sense, contextualisation improves the usefulness of statistical methods.

The clustering approach used was developed from the typical automatic clustering of Salton's document indexing algorithms (Salton & McGill, 1983). For the most part, these measure the differences between documents using Salton's TF:IDF (Term Frequency: Inverse Document Frequency) measurement. This algorithm is dependent on the whole corpus you are working with and the results must be recalculated every time you add a new document, like ID3. The advantage of using Salton's approach, over that of Quillian, is that this document 'distance' measure can subsequently be used in retrieval. This algorithm is reflexive in nature as it normalises the distance measurement against the length of the documents and the term frequency against the total term frequency in the corpus. When attempting document retrieval on a database with this method words that occur frequently in the corpus will give each document a low score, words that occur infrequently in the corpus will give those documents that contain them a high score. This algorithm can then be used compare different documents in a corpus along axes defined by words they have in common.

Taking this as a starting point it is necessary to consider that we are more interested in set membership than normalised comparisons. We are looking to cluster together those threads which discuss similar topics. These threads will be of varying lengths and each will discuss the topic to different levels of detail. A simple heuristic would indicate that the longer the thread the more detailed the discussion. This heuristic can be inferred from the observation that a longer thread usually involves messages asking for clarification on specific points, and more messages from tutors giving a guiding hand. The content of these discussions is, in the majority, going to be more useful to students in other years, and other threads that discuss the same or related topics should be linked to it. So the selection of a distance metric which places a short discussion on a topic as being close to a large discussion on the same topic, but not vice versa, is not unreasonable. This approach is not without precedent and a similar approach has been used in previous work on the generation of models of a user's topic interest for filtering (Masterton, 1995). This is related to the use of key word sets as user profiles in recommender systems (Resnick & Varian, 1997). The methods used by all these systems have very similar backgrounds, most having descended from the work by Salton and McGill (1983), and you could equally well choose to use any one of them instead of the one given here. The reasoning for developing my own was taken so I could have full control over the process used.

*Developing the clustering algorithm*

Automatic clustering was reapplied to TMA 1, to improve the results of the manual clustering, and was the only technique applied to TMA 2. Here I will describe the specific algorithm which was used. Starting with the individual threads for each of the years the words contained in the stoplist (appendix 2) were removed. Each thread was indexed by linking each term (i.e. each individual word) to each thread in which it occurred. This generated a database of words, listing for each word which threads contained it. The database was then used to create an index file listing each thread and the number of words it contained.

This process then enabled the comparison of files (threads) to determine how similar they were. When comparing file1 and file2 we would look at every word in the database that is in file1 and file2. For each word in turn we then work out its indexing value for each of file1 and file2 given by:

$$\left(\frac{\text{Number of occurrences of word in file}}{\text{Length of file}}\right) * \text{Log}\left(\frac{\text{Total number of files}}{\text{Number of files in which this word occurs}}\right)$$

Equation (1): Term frequency multiplied by document frequency.

The left term is the 'term frequency', *tf*, and the right hand term is the 'document frequency', *df*. By calculating this value for each file and then taking the absolute difference between them we get a measure of how similar the files are for this term. Because this function returns a Term Frequency/Inverse Document Frequency measure (Salton & McGill, 1983) the log function means that the fraction (number of occurrences/length) is made exponentially smaller depending on how many files it occurs in, i.e. words which are common throughout the corpus are of little value in distinguishing between documents in that corpus. This function returns values of no more than 1; higher values are more indicative that this word is highly occurring in this file and very relevant. The scores give a way of comparing documents along this dimension. We want to know if our documents are similar for this word so we take the absolute difference between the two values. We then sum these for every word occurring in both files. We then multiply this final sum by (total number of words in *both* file1 and file2 / total number of words in file1) giving a final equation of:

$$\sum_{\text{All words}} \text{ABS}\left(\text{TF-DF-File1} - \text{TF-DF-File2}\right) * \left(\frac{\text{Total number of words in files 1 and 2}}{\text{Total number of words in file 1}}\right)$$

Equation (2): The first comparison function used to create clusters.

This result gives us a measure of how close file2 is to file1. This similarity measure is not reflexive, but as stated earlier, in the example on sets, this is not what we are looking for.

The comparison function, equation (2), was revised because the original *df* term in equation (1) used a logarithm in the function meaning that it varied between 0 and 2. This term affects the distance measure dependent on the overall frequency of a specific word in the corpus as a whole. By altering the *df* term to be the one following (in equation 3) it now varies between 0 and 1. With the logarithmic functions words that occurred highly infrequently in the corpus as a whole would have a large impact on the distance metric. The choice of the exponential function flattens out the effect of highly infrequent words on the distance metric, while ensuring that the measure has a maximum value of 1:

$$\left(1 - e^{\left(\left(1 - \left(\frac{\text{Total number of files}}{\text{Number of files in which this word occures}}\right)\right)/2\right)}\right)$$

Equation (3): Replacement for the df term in equation (1).

The effect of this choice can be see in figure 3.5.

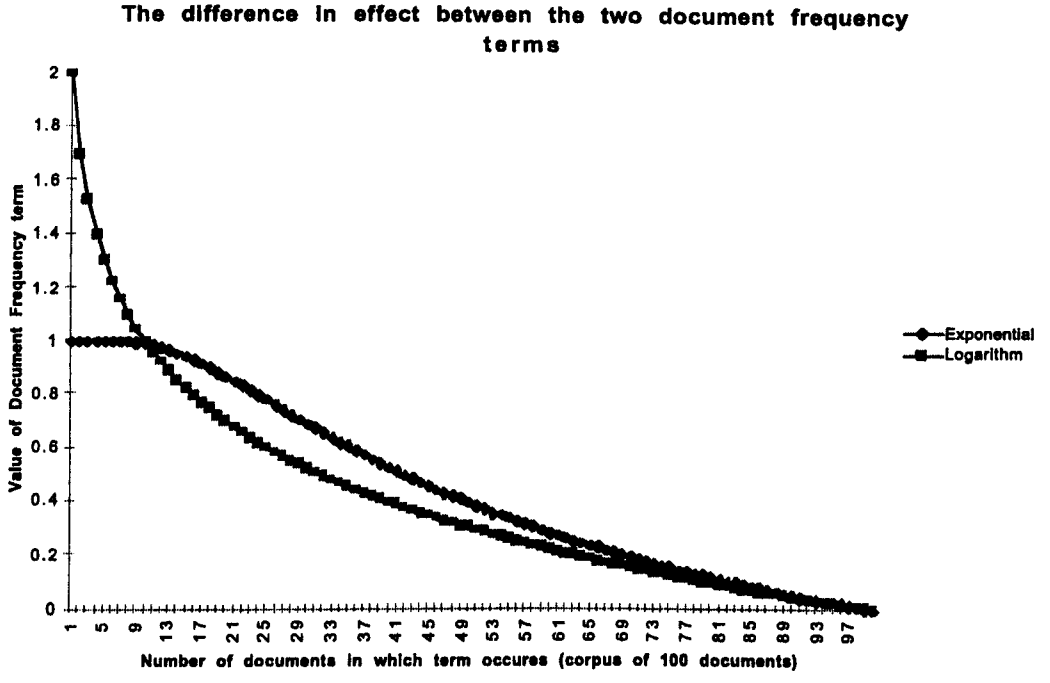


Figure 3.5: The difference between using the exponential and logarithmic terms for the document frequency measure.

Secondly the *tf* term used in equation (1) has been adjusted to take into account how similar the number of occurrences of the term in question is in each document being compared. First we work out *d*, the 'difference' between the *tf*df* for each document using the new term in equation (3), the result of this could be negative. The more different the distributions are the more suppression is put on the distance metric by using the form $(1 - d^2)$, where the d term has been multiplied by itself to ensure it is always positive. Now, rather than working out the absolute difference, we instead calculate the 'average', *a*, value of the *tf*df* for each document. This is then summed for every word occurring in both using the equation $(1 - d^2) * a$. This then replaces the first term in equation (2) to give equation (6). This is a reflexive measurement.

The difference, *d*, is simply:

$$\left\{ \left(\frac{\text{Number of occurrences of word in file1}}{\text{Length of file1}}\right) \cdot \left(1 - e^{\left(\left(1 - \left(\frac{\text{Total number of files}}{\text{Number of files in which this word occurs}}\right)\right)/2\right)}\right) \right\} - \left\{ \left(\frac{\text{Number of occurrences of word in file2}}{\text{Length of file2}}\right) \cdot \left(1 - e^{\left(\left(1 - \left(\frac{\text{Total number of files}}{\text{Number of files in which this word occurs}}\right)\right)/2\right)}\right) \right\}$$

Equation (4): The difference between the *tf*df* terms for each document.

The average, *a*, is:

$$\frac{\left\{ \left(\frac{\text{Number of occurrences of word in file1}}{\text{Length of file1}}\right) \cdot \left(1 - e^{\left(\left(1 - \left(\frac{\text{Total number of files}}{\text{Number of files in which this word occurs}}\right)\right)/2\right)}\right) \right\} + \left\{ \left(\frac{\text{Number of occurrences of word in file2}}{\text{Length of file2}}\right) \cdot \left(1 - e^{\left(\left(1 - \left(\frac{\text{Total number of files}}{\text{Number of files in which this word occurs}}\right)\right)/2\right)}\right) \right\}}{2}$$

Equation (5): The average of the *tf*df* terms for each document.

Giving finally:

$$\sum_{\text{All words}} \left( (1 - d * d) * a \right) * \left( \frac{\text{Total number of words in files 1 and 2}}{\text{Total number of words in file 1}} \right)$$

Equation (6): Replacement for equation (2) and used in comparing threads.

Adjusting the power to which d is raised in $(1 - d^2)$ alters how much effect the difference between two *tf df* factors has. Both *d's* are raised to the same power, it is multiplied by itself to guarantee the sign will always be positive. Making the comparison dependent on the length of one of the documents means that this function is not reflexive, which is what we wanted. So the effect of this function is that the more common a word in the corpus as a whole, then the less influence it has (the *df* term). And the more similar the ratios of number of occurrences to length of file for each document then the more influence it has (the $(1 - d^2)$ term). Finally if file 1 is shorter than file 2 then the larger the measure, and vice versa the smaller the measure, hence the lack of reflexivity and the fact that larger clusters subsume smaller ones.

*Applying the clustering algorithm*

Once these similarity measures have been calculated the next task is to cluster the threads together. There are two approaches to this. The first is to start by treating each thread as a single cluster, starting with a low threshold. You then find all the clusters which are closer than this threshold, and combine these into a new cluster. By repeating this process and slowly increasing the threshold you eventually get one big cluster. By looking at the number of clusters at each threshold value there will be plateaux where the number of clusters is constant or is decreasing very slowly. These areas of stability provide useful clustering. The alternative approach is the direct opposite. Start off with all the threads in one cluster and slowly reduce the distance threshold. The second approach is better in this case because the first isolated threads to leave the main cluster are generally those of little or no use. Later on, the bigger clusters break up into several clusters containing at least 2, but often 3, 4 or a larger number of threads. The majority of these provide useful data. By repeating this process for each year, and for each combination of two years, and for all years together, clear patterns emerge. You see strong clusters from one year appearing in the all year clustering, and the same pattern with cross year clusters as well.

Figure 3.6 shows the effect of starting at a low threshold value (step one) and slowly increasing threshold value with each step. In this figure the data has been normalised with the y-axis representing the number of clusters divided by total number of possible individual clusters. The five lines show the number of clusters formed for each year individually (94, 95, 96), the average of these values (94+95+96), and the effect when we put all the threads together and clustered (94/95/96). A value of 1 on this scale means that all threads are in their own cluster.

These charts clearly show the plateaux in the clustering process. Also, near the start you can see a quite rapid initial break-up before the first plateau, due to the smaller, irrelevant, threads breaking off. The lines showing all three years combined, and the numbers of clusters you get at each step if you combine all years do not clearly show the complexity here, but it is important to consider that although the number of clusters is quite similar, the membership and sizes of these clusters is not the same.

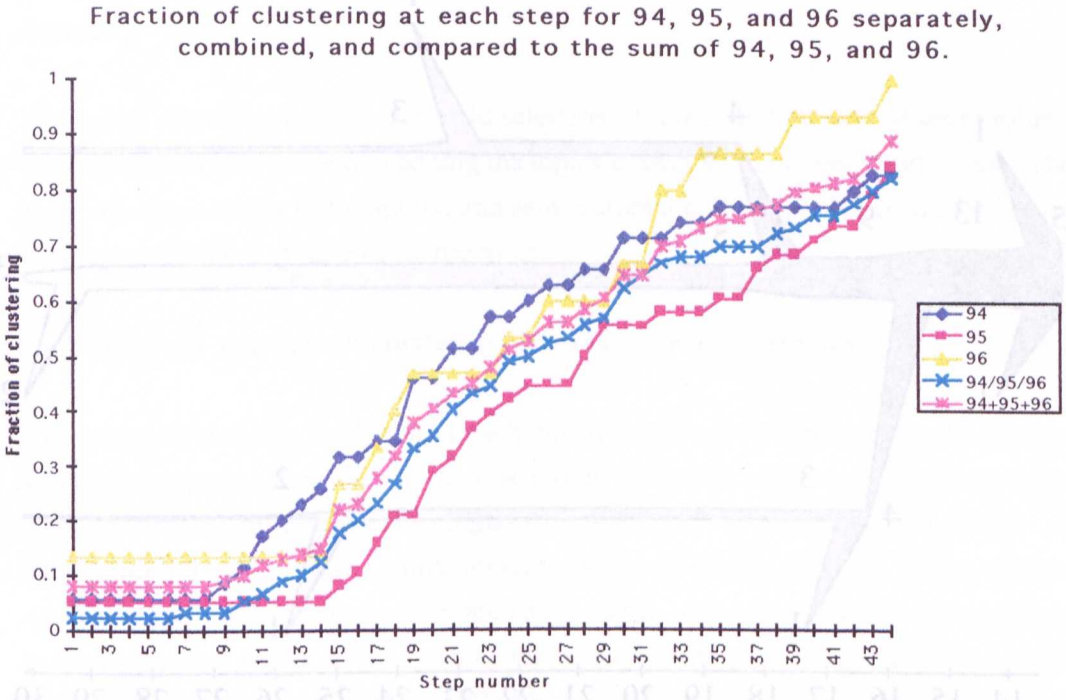


Figure 3.6: This chart shows the fraction of clustering for each step in an increasing threshold for TMA2 in 94, 95, and 96, individually, the average of these values 94+95+96, and when all the threads were put together and clustered 94/95/96.

Although there is no clear quantitative measure for when you have a good set of clusters, there are a few rules of thumb that come from this initial assessment of clustering. These have proven effective for this course's data, but different discussion forums and domains may need different or adjusted measures. The main rule, traditional in clustering, is look for plateaux. We also found that the threshold values around the points where these occur are reasonably consistent, at least for this corpus. We find that the number of clusters is usually about 1/2 - 2/3 the total number of threads. This doesn't sound very good but then a more detailed look at the threads themselves reveals, as reviewed in section 3.2.1, that there are a lot of short threads where discussion didn't take off or which are simply information messages from tutors. Often these are threads of less than 5 messages. This effect can be seen more clearly in Figure 3.7 which looks at the clustering of 1996 TMA2 threads between steps 14 and 30. There are clear sections, between steps 19-23 and steps 26-29, where no break-up occurs (i.e. plateaux). Between steps 24 and 30 you see 7, 8, and 9 clusters (out of 15 total threads) i.e. the number of clusters is approximately 1/2 - 2/3 the total number of threads. It is in this range that the choice of when to stop should be made.

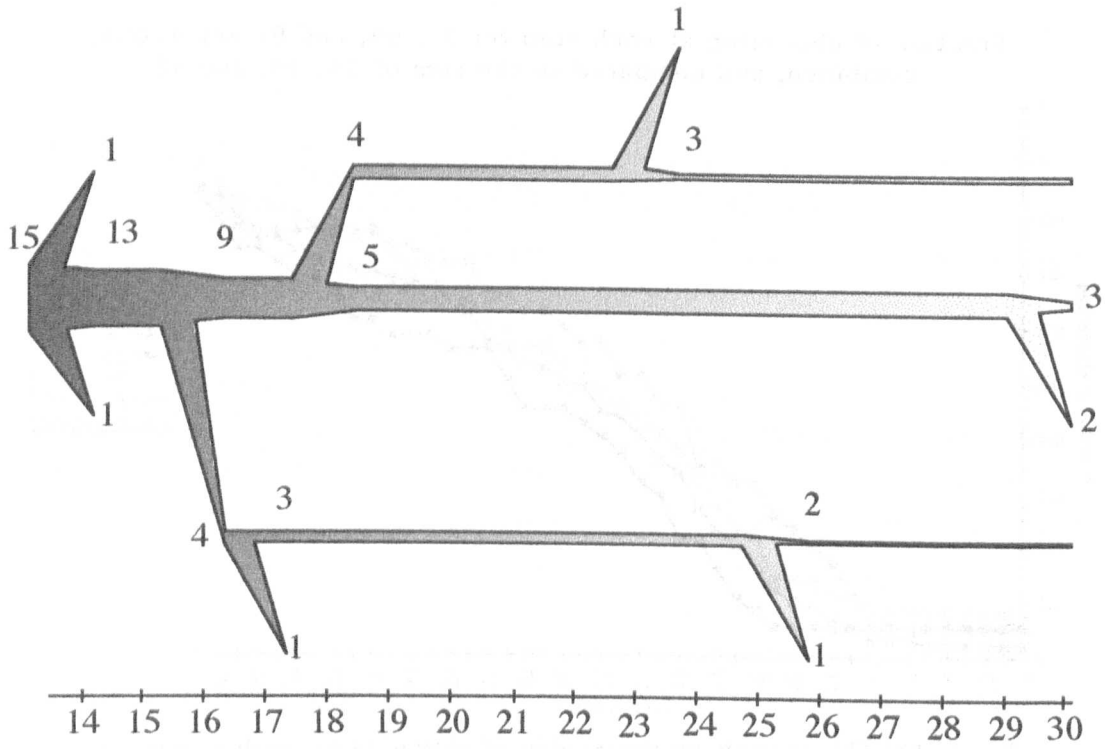


Figure 3.7: Cluster break up for 1996 TMA2 threads. The numbers indicate the number of threads in each cluster. The steps move from left to right and represent steps 14 to 30 from figure 3.6. Those clusters containing only 1 thread are not continued.

The clustering tends to rapidly drop out the trivial threads, though across years it will create clusters of 'welcome to this conference' and 'I'm having a problem with the conferencing system' messages. Based on these observations, the best approach is to remove all threads which form their own clusters and contain only 1 or 2 messages. From then on you really have to go for manual editing, which will be discussed in section 3.2.4.

Applying these techniques to the data for past years of TMA 3 confirmed, as expected, that much lower thresholds were necessary for clustering. The MBA course is very practically based and by the time the students have reached the third block they are concentrating on their own problems which have diverged so much from the common ones of earlier blocks that there is much less that can be done to support them, at least in terms of vicarious learning. At this point they are engaging in collaborative problem solving on their own specific problems. It is perhaps now that a brainstorming agent, as reviewed in chapter 2, would prove more useful.

*Summary*

The use of clustering enables the rapid selection of topics and threads on those topics. This then assists the process of editing the topics down into their core information. The approach also enables us to rapidly, and semi-automatically, identify and discard those messages which are of little or no use to us.

### 3.2.4 Manual editing of clustered threads to create stories

As mentioned earlier in section 3.2.1 each thread contains a certain number of messages which do not really contribute to the discussion. These messages need to be removed from the threads to increase the 'signal to noise ratio'. To this end each cluster of threads was edited to create a 'story' about the topic which they discuss. These stories, as already mentioned in section 1.3, are structured in such a way that the first message is about a student with a problem which relates to the topic the VP has identified in the discussion. The students are then able to ask questions to get at other 'cases' which elicits a response from the VP providing more details which may come from the same thread of discussion as the first message, or other related threads identified by the above clustering process. An example of this is shown later in section 4.2, figures 4.3 and 4.4.

The structure of a complete story is made up of a number of related 'cases' which give information about various aspects of the topic. The details of how the cases are accessed by the system will be discussed later in part B of this chapter. Here we describe the case-base itself and how the threads, concordance, and the clusters help feed into this.

Figure 3.8 shows how the case-base is organised. It consists of four types of objects: the basic messages make up threads of conversation, threads which discuss the same thing make up topics and topics which discuss similar things make up stories. With the semi-automatic methods described in the previous section we had reached the point of topics, now an element of manual editing was necessary to complete the case base. I then chose which topics to merge into stories and which to separate, if any, although this was rarely necessary. The final case base was presented to the tutors to give them a chance to request any changes and to get their approval; unfortunately they did not provide any input into the structuring process itself.

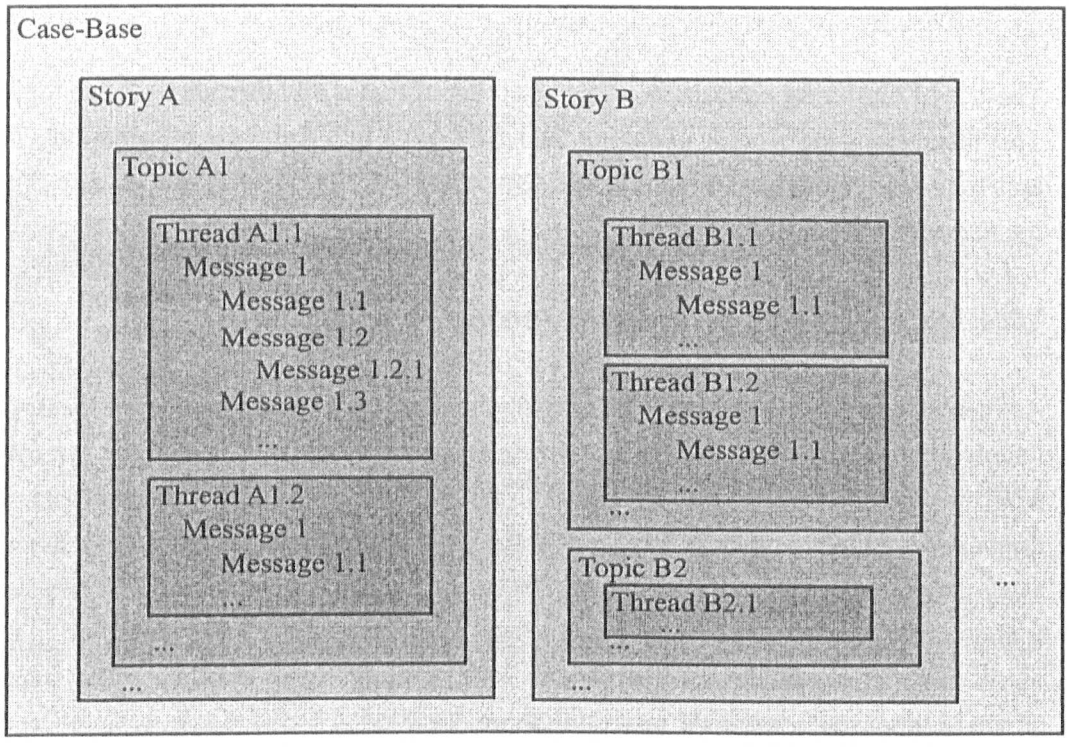


Figure 3.8: The composition of the case-base, messages, threads, topics, and stories.

*Summary*

This section has given an overview of how the case base of topics and stories was created. The four step process we have gone through has been assisted by a number of computer programs, a summary is shown in Table 3.1.

| Step | Process | System | Section |
|---|---|---|---|
| 1 | Group messages into threads | Threading program | 3.2.1 |
| 2 | Identify keywords & phrases for indexing | Concordance program | 3.2.2 |
| 3 | Group threads into topics | Clustering program | 3.2.3 |
| 4 | Generate stories | Manual editing using human judgement on output from the above programs. | 3.2.4 |

Table 3.1: Process and systems used in the development of the case-base.

Inevitably there are some points at which human involvement is necessary, especially when it comes to deciding what to use in the cases, and what to discard. The more input by the tutor at this stage the greater the likelihood of the cases being of use to future students. The tutor is also able to quickly confirm the quality and applicability of the content. The effort involved here is that of an 'editor' who needs to remove extraneous material without changing the core of the information - no rewriting is involved. A domain expert, like a tutor, can add extra value at this stage. However the stories used in these experiments were generated by myself and cursorily examined by the tutors. Finally, although this provides a sound case-base structure there are still many issues that remain to be resolved. It is now up to the implementation to use this case base for the benefit of the students.

## Part B: Contextualising Frequently Asked Questions

We will now explore how the case-base derived using the techniques in Part A can be used in practice. We have seen how to index and retrieve cases, and we have also found that there are common discussions that occur across years.

In section 3.1 when we first looked at FAQs, we identified context as the key problem in their use, users just didn't get answers when they needed them. The Vicarious Learner (McKendree, Stenning, Mayes, Lee & Cox, 1997) overcame this problem by linking FAQs to on-line course material. In this way the context is provided by the materials. The Creative Management course does not have on-line course materials, so this approach is not available. However, the cases were generated from past conference discussions, and the course was to continue to use conferencing. ContactFinder (Krulwich & Burkey, 1996) worked proactively in electronic conferences, identifying people seeking information and trying to link them to contacts who might be able to help. We are looking for a similar approach where the system will monitor the conferences looking for topics of discussion which it has in its database. Once identified the system can then link the past information to the current discussions. Such a system would function as just another user in the conferencing structure. Hence we coined the name 'Virtual Participant' to describe the system.

The type of conferencing we are working with is asynchronous and text-based. One problem of asynchronous communication is that tutors are not always available and messages take time to get answered. An approach that has been taken to help students get over the "first public message" barrier and encourage them to stay involved in the conferencing is to welcome them each personally the moment they first logon and set them a simple task (Salmon & Giles, 1997). This simple task asks the student to post their first message to the group saying who they are and why they are there - a standard ice-breaker. In an asynchronous conferencing environment greeting new students is an ideal task to be handled automatically. The Virtual Participant could handle the automated welcoming messages, and any other similar messages. This would enable at least some support to be available synchronously.

The approach taken with the VP is similar to that used in the 'Bugtales' project (Eisenstadt, 1993; Eisenstadt, 1997), where you would tell the system, by e-mail, your problem as a story and it would then send you back a possible solution as another story. The approach taken with the Virtual Participant was that it observes the story telling, as part of the group. As the story is told it is reminded of one of its own

stories, which it then passes on to the group. Part A has provided us with the raw materials necessary for this approach. Part B will now look at how the Virtual Participant functioned as a program. The next chapter then looks in more detail at how the first prototype of the VP, which became known as 'Uncle Bulgaria', fared with the students.

## 3.3 Technicalities of the implementation

At the beginning of this project there was a lot of interest in the possibilities it raised. The major stumbling block was that no one really had any ideas about what would prove to be effective. So an initial prototype was necessary to give the tutors and students something to comment on. Other projects and research reviewed in chapter 2 gave some ideas as to what might prove effective. Certainly there was a clear necessity of contextualising the topics in the case-base so that the students would receive them at a time when it was most useful for them. However, the important or trivial aspects of such a system and its interaction with users was not clear.

Having opted for using conferencing to interact with the students, it was then necessary to interface to the FirstClass conferencing system itself. The software was developed on the Apple Macintosh platform, partly because FirstClass was originally a Macintosh program and best supported on that platform, and Macintosh Common Lisp provided good network support. In addition the Macintosh family of computers supports Apple Script and Apple Events which allow control of one application from another. Unfortunately there was only minimal support for this with the First Class (FC) client. The First Class server does support a command line user interface (CLUI) in the style of a vt100, although in a rather ad-hoc manner. This gave the system one way to link to the server. Using the CLUI the software could gain access to almost all the functionality within the 2.6 version of the client. This interface is shown in figures 3.9, 3.10, and 3.11. Using this interface the software was able to read messages, send messages, move through the hierarchy of conferences, and check message histories.

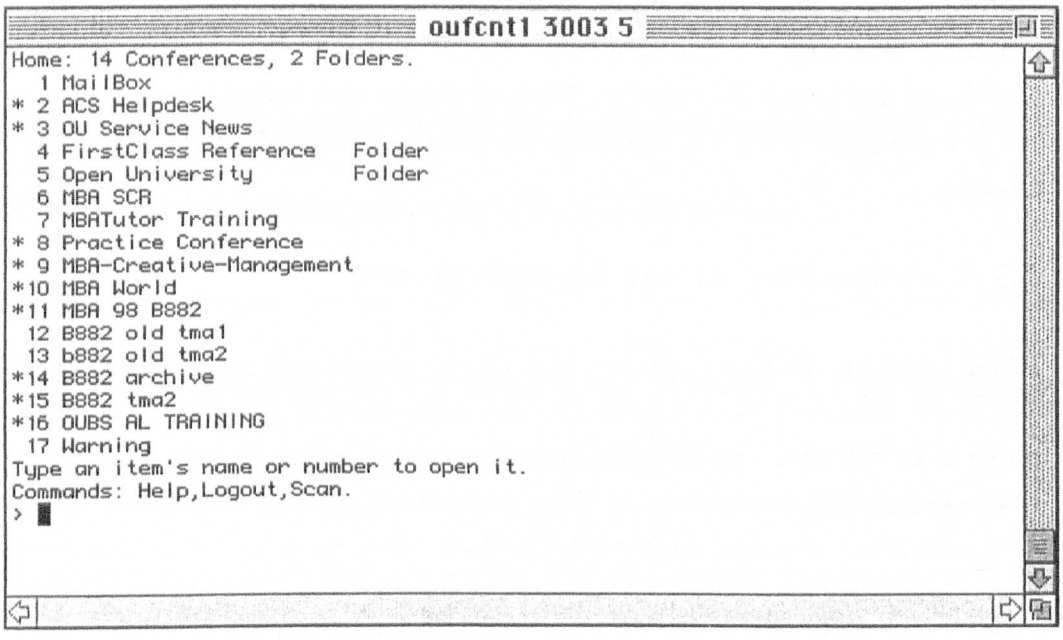


Figure 3.9: Example of FirstClass CLUI showing the desktop conferences. Stars indicate unread messages.

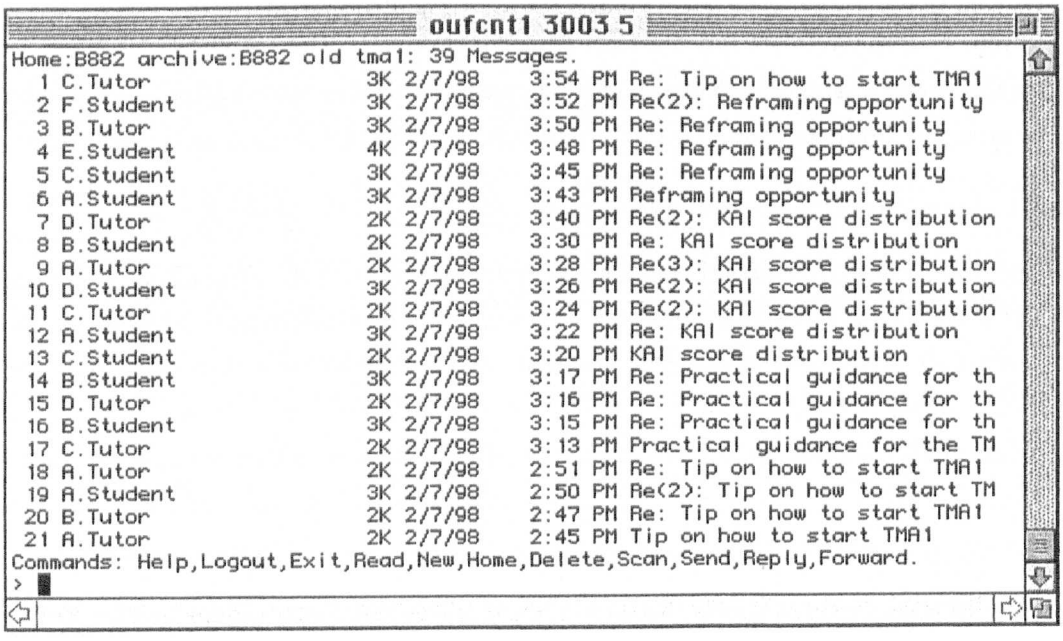


Figure 3.10: Example of FirstClass CLUI showing messages within a conference.

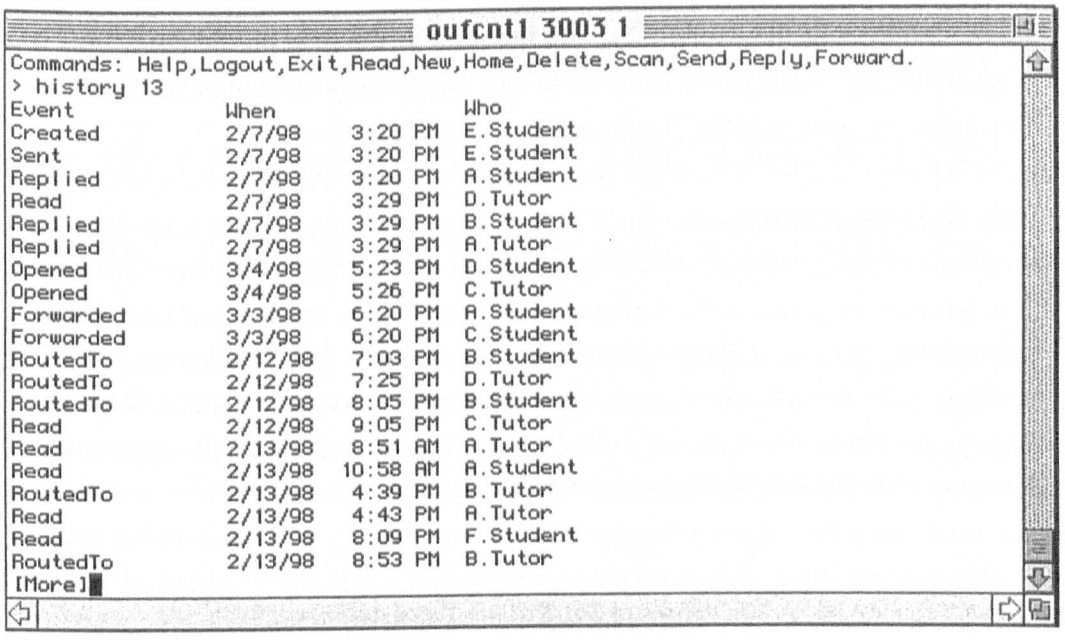


Figure 3.11: Example of FirstClass CLUI showing message history.

The 'history' function available in FC is a powerful (and almost unique amongst conferencing systems) feature of FirstClass. The server keeps track of all the users which have read, forwarded or replied to every message. This record forms a history, like that shown in figure 3.11. This facility comes in very useful for keeping track of what has been going on within the discussions and makes it possible to measure the level of participation reasonably accurately. However like many things considered to be 'big brotherish', students subvert it, some logon every few weeks and open all the messages 'just in case their tutor is keeping track', and others have been able to read the messages without entries appearing in the history list. At the time First Class, unfortunately, did not have the ability to thread messages[2], and this sometimes made it hard for students to keep track of discussions. However by using the subject fields and the history information it was possible for the prototype Virtual Participant to keep track of the threads.

Finally, to interact with FirstClass and with the students it was necessary to have a user ID to represent the VP to the students. Mindful of the need to avoid over anthropomorphising (Koda, 1996) the system, and looking to pick something relevant to the course, I chose the name Uncle Bulgaria (Beresford, 1968). Great Uncle Bulgaria was the patriarch of the Wombles who lived on Wimbledon Common in

[2]Version 5 of FirstClass which became available after the end of the study did include this feature.

London and spent their time recycling the rubbish dumped on the common. It was intended for the name Uncle Bulgaria to be a metaphor (something commonly used in B882) for the system, in that it behaved as a 'knowledge recycler'.

## 3.4 System rationale

This section uses various components from the KADS system analysis and design methodology to give a clearer picture of the behaviour and functionality of the first prototype. We begin this by presenting a system overview document. This forms a "comprehensive and concise summary of the whole system's objectives and functionality" (Tansley & Hayball, 1993).

A system overview document consists of four sections and is intended to provide a top-level overview of a system and its interactions with its (possibly several) users.

*Objectives of Prospective System*

This section is not intended to specify the higher level research goals of the project, but rather to specify more clearly the precise objectives of the first prototype and its users.

First let us consider the different "agents" within this system.

- Tutors
  - Normal course tutors
  - Electronic tutors
- Students (could be broken down along many axes)
- The course team
- The Virtual Participant
- Myself

The concerns and wishes of these agents are not necessarily congruent. The students wish to pass the course, the tutors wish the students to pass the course and get paid, the course team want the students to pass the course, get paid, and look good, I want to see what happens when the VP is introduced, and so on.

This immediately raised two clear aims for the system.

- It must not interfere with the tutors and students in achieving their goals, so it must be unobtrusive and easy to ignore.
- The system needs to be introduced slowly so as not to disrupt normal teaching patterns. To this end I would like to observe a few interactions between the system

and the students, but I do not need an excessive amount. This allows the assessment of the system while keeping within the bounds of the first aim.

What emerged was that the first prototype not only had a proof of concept role, but also a political one. The aim of the exercise was to win over the hearts and minds of those whose environment it impinged on the most - primarily the electronic tutors, but also the students.

The only risk of the approach was that there might not be sufficient interaction for a meaningful study. However it seemed likely that the majority of the key problems would show themselves even with limited interactions. Monitoring the progress of the system would help here. If the data was insufficient it would be possible to make changes midway though the course. It was suggested that if a prototype was not available on time then a 'Wizard of Oz study' could be conducted to gain similar insights.

To summarise, the high level aims were to convince the tutors and to see what happened. Possibly these were not clear research objectives at this preliminary stage, but in this area there was no clear idea of what would happen, what students and tutors would think, or how they would use it. Only with a slightly clearer picture of this could the objectives and the prototype be revised and improved.

*System Functions*

This section of the KADS methodology is quite straightforward and is intended to be a "list and summary description of the complete set of functions that the system will support" (Tansley & Hayball, 1993). It was intended that the system would perform the following functions:

- Recording messages and history of messages from FirstClass.
- Storing a database of past cases.
- Storing the current state of the system.
- Matching current FC discussions to past cases.
- Posting of 'matches' with past cases to FC.
- Identifying, and dealing with, public requests for more information.

*Provisional System Structure*

This section provides a high-level schematic representation of the system and its component sub-systems. This schematic representation of the provisional system structure can be seen in figure 3.12. There was no change from this in the implemented first prototype structure. Later on when we come to look at the functioning of the system we will see how data and tasks flowed between the different components.

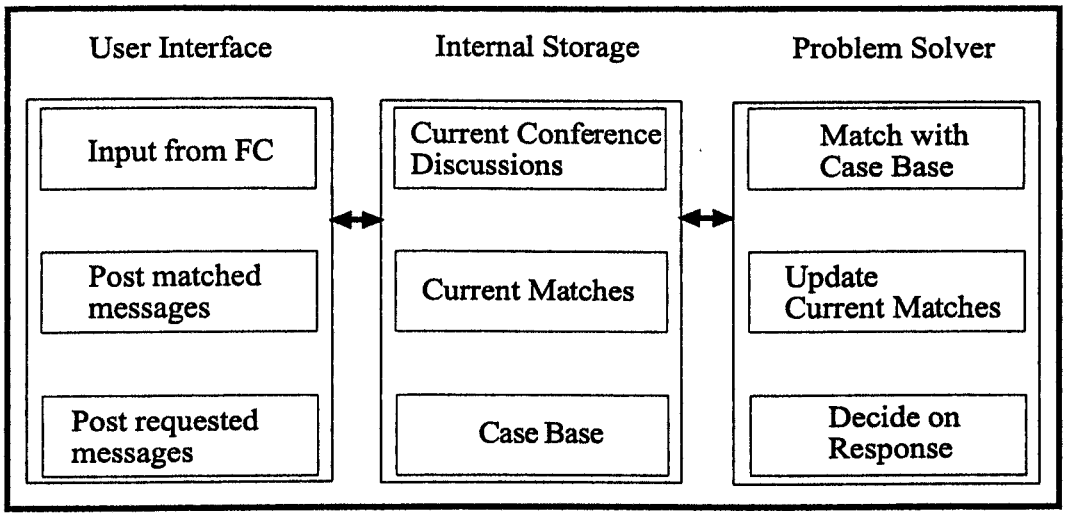


Figure 3.12: Schematic diagram of provisional system structure.

*Provisional Information Requirements*

In the case-base we store (as Figure 3.8):

- Anonymised messages from previous discussions.
- These are arranged as threads.
- Threads are clustered into a topics.
- Topics have been grouped into stories by manual editing.
- A data structure representing each of the stories, the matching threshold, and a current match level of 0.

In the thread database we store:

- Messages posted to FC.
- These are arranged as threads.

- The 'who has replied to this message' history information for each message.
- For each message in the thread a value, updated from the message it replies to, measuring how closely the message matches each of the stories in the case-base.

For the last point the update is made by taking the current value from the previous message and adding to it for each and every additional keyword or phrase matched. Those messages with matching values greater than the given threshold trigger the story.

This concludes the system overview document. In the next section we will now look at the data flow within the first prototype and in the section after that at the generic task model.

*Process analysis: Data flows*

The following data flow diagram shows the process that the VP goes through in dealing with new messages. Input is drawn from the 'History' information and the 'IN' box, output goes to the 'OUT' box. These refer to the functions in FirstClass accessed through the CLUI. The information gained from the history, and the messages themselves, are stored in the 'Threads' database.

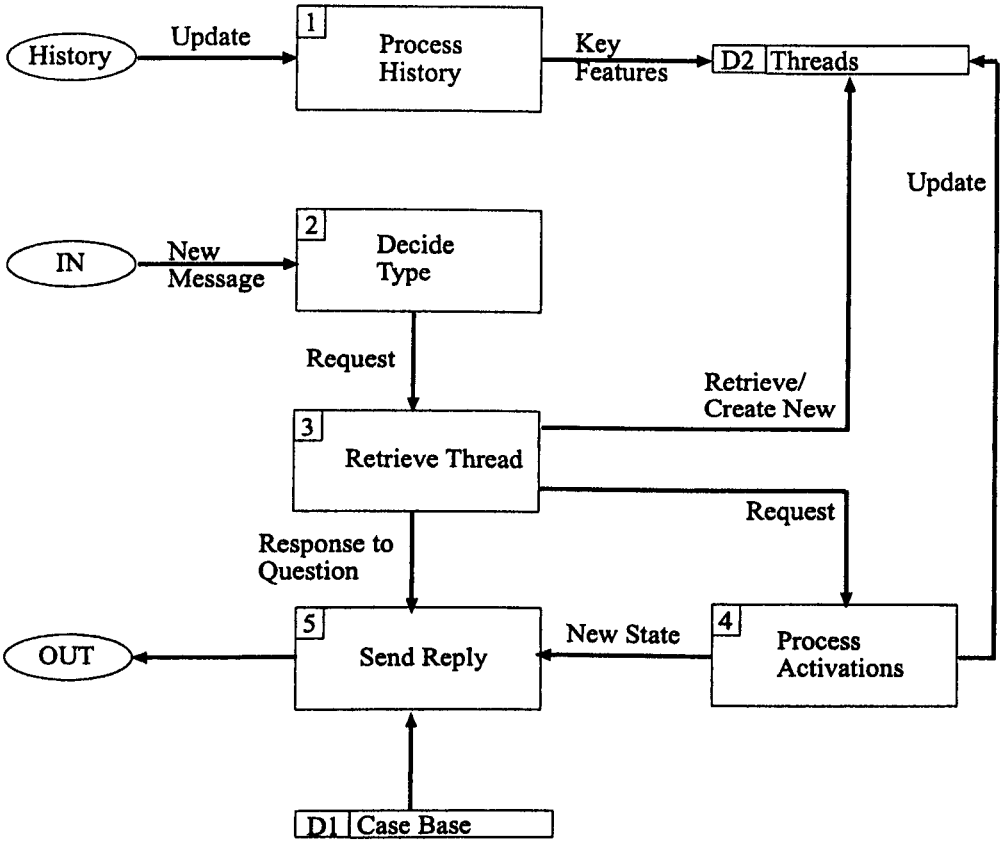


Figure 3.13: Data flow diagram for First Prototype.

The data flow diagram presented in figure 3.13 provides a schematic of the data flows involved in the VP. These processes represent the following:

1. Once the VP has logged in to FC through the CLUI the first thing it does is download the new history information. This is processed in an oldest to newest fashion, separating the key information of who created a message and when, from records of who has replied to a message and when. This information is added to the details of the associated thread.
2. When the history information has been processed any incoming messages are then retrieved, again in an oldest to newest fashion. At this point the message is classified as 'new' or a 'reply' to assist in processing. The message is then passed on as a request for processing.
3. From the message information the appropriate thread is retrieved or a new one generated. If the message (which must be a reply) is identified as containing a question this is passed to process 5. Otherwise all messages with their identified threads are passed on to the next process.

4. The matching process identifies any case in the case base which matches with the thread above its threshold value. This is then passed to process 5. Either way the thread database is updated with the new state.
5. Finally the appropriate message (if any) is retrieved from the case base and passed to the outbox in FirstClass.

This section has shown the data flows involved in the first prototype Virtual Participant. The next section looks in more detail at the selected Generic Task Model and its relationship to the function of the first prototype.

*The Generic Task Model.*

From the generic task model (GTM) library in (Tansley & Hayball, 1993) the task of the VP is a "System Analysis: Identification" task. The GTMs in this library which could be used were "Systematic Diagnosis", "Systematic Refinement", "Correlation - Assessment", and "Monitoring". The first two of these are very similar, and the second two are also closely related. I finally selected the "Correlation - Assessment" GTM as being the best model.

The Inference Structure used by "Correlation - Assessment" is very similar to that used by a generic help desk system, for example, where you might be identifying faults and looking for possible solutions to them. The input/output descriptions we would see for the VP would look like Figure 3.14.

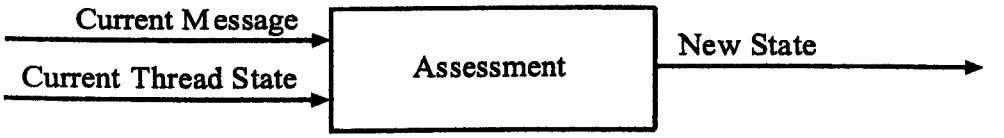


Figure 3.14: Input/output descriptions for the VP.

The current message coming in to the assessment system is a message that has been received from FirstClass and is awaiting processing. The current thread state is the set of values for matches between the case-base and the current thread at this point in the thread. If the message starts a new thread the current value of all matches with stories is set to 0. The new state produced is associated with the current message in its thread.

To make the matching process clearer consider the analogy of an inheritance hierarchy. Each message inherits its current state from the node above. The state is then modified with respect to the current message. Every daughter message then inherits the state of the parent message. This allows each branch develop into a new topic area.

We have adapted the basic generic task model for Correlation - Assessment in a straightforward way to reflect the inference processes going on within Uncle Bulgaria when dealing with messages. This is shown in figure 3.15.

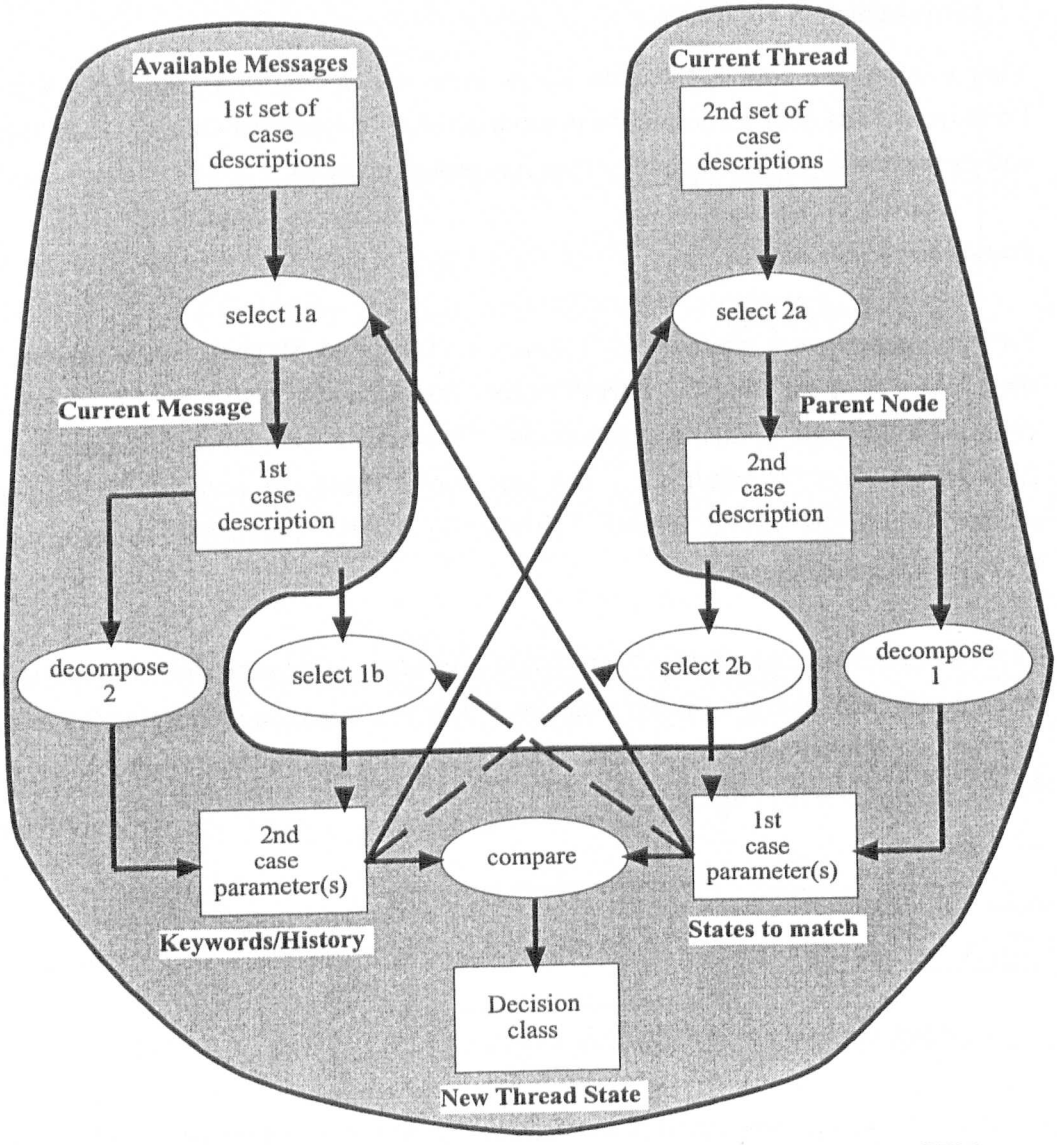


Figure 3.15. The Task Model for Uncle Bulgaria - based on the GTM Correlation - Assessment (Tansley & Hayball, 1993).

The original GTM is shown, with those parts used by the Uncle Bulgaria system shown within the grey area. The new task model shows the various stages of the system's function. The incoming messages are used to select the correct thread, this may also involve generating a new thread. The messages are then processed one at a time. When an incoming message has been selected as the current message its contents are decomposed into keywords and the corresponding history data has reply

information extracted from it. The history data is used to select the correct parent node in the thread, which then has the current match state extracted from it. The current match state and the keyword information are fed through the comparison stage to produce a new set of states for this new node in the thread.

This section has given an overview of the processes performed by the software. The goals of the first prototype have as much to do with convincing the people involved as they do for proving the concept. The second prototype proved a much more significant implementation and I will come back to this kind of analysis later for the second prototype, in chapter six, to highlight the important differences and to discuss areas for improvement and future work.

## 3.5 Summary

This chapter has shown how the initial ideas and background information grew into an initial prototype. We have looked in detail at the content of previous year's discussions and seen how they can be semi-automatically distilled into key topic areas for the course. These topics make the foundation of the case base. We then moved on to look at how to make these topics available to the students. The cases are delivered by a mechanism we call the Virtual Participant. This is an agent whose role is to read the discussions within the conference and as these topics occur again to use that recurrence to contextualise the stored cases. The next chapter looks at what happened when the system was introduced, first to the tutors and then to the students as the course presentation unfolded. Finally, we look at the use the students and tutors made of the system and what they thought of it.

# Chapter 4: Tutors, Students, and Virtual Participants: Experiences with the First Prototype

## 4.0 Preamble

This chapter looks at what happened with the first prototype system, while the course was running. We start by briefly reviewing the original aims of the system and some of the important assumptions that were made. After that we move on to look at how the tutors felt when the system was introduced to them. Then we look at the interactions between the students and the first prototype which occurred during the course. In the next chapter we will review the feedback from the students and how this enabled us to revise and refine the system, and develop the second prototype.

## 4.1 Aims of the Virtual Participant

The original aim of the Virtual Participant was for it to function entirely autonomously, answering student questions, and even automatically updating its knowledge base. This was an idealised solution typical of some AI approaches. This idealised approach missed the key goal of supporting tutors and students in electronic conferencing; by not supporting their needs the system lacked any connection with its users. The developments we have made have been aimed at addressing this smaller goal.

In previous chapters we have developed a way of approaching this key goal. Initially we identified a teaching resource, our knowledge base of previous years' students' experiences. We then went on to show that the information contained in this knowledge base is of most use to the student when contextualised within their current situation. So we developed a mechanism to identify context, which became the core of the Virtual Participant. Perhaps this was not the ideal of a fully functioning autonomous agent with which you could have a discussion but it did begin to address the key goal of supporting students and tutors. Now it was necessary to assess the effectiveness of our approach.

It is necessary to find out how the Virtual Participant affected the conferencing environment, and what the tutors' and students' perceptions of it were. If we could observe reactions to the VP they could then be used to change the design to increase its effectiveness. Perceptions can be assessed through questionnaires and other, more

personal, ways of seeking feedback. Ideally, we would hope to observe positive effects like the following:

- Increased student participation and contribution to discussion.
- Students reading messages from the Virtual Participant.
- Students asking the Virtual Participant questions.
- Tutors following up the Virtual Participant's messages.
- Discussions about the Virtual Participant and users' opinions of it.
- Feedback from the tutors about the students', and their own, opinions.

Alternatively, we might observe no change in some of these aspects, or even a negative effect. If we were to observe a decrease in participation this would be a serious problem for the approach. On the other hand no change in levels of participation would be considered reasonable. At this point we were just trying to map out the territory. We did not know what would help and what would be effective, which is why this assessment of the first prototype Virtual Participant was so important. By observing what the students and tutors thought of it and how they reacted to it we could begin to fill the gaps in our knowledge and feed this back into better designs.

## 4.2 Introducing Uncle Bulgaria

For the Virtual Participant to interact with the students it was necessary for it to have a name, a user-id for First Class, so that it could be identified within the conferences. After some discussion we (the B882 tutors and myself) selected the name of the Womble 'Uncle Bulgaria' (abbreviated to UB). The metaphor we were aiming for was that the Virtual Participant would recycle the useful discussions 'left behind' by previous students.

When I introduced the idea of calling it Uncle Bulgaria a number of the tutors thought that this was a fantastic metaphor (metaphors are used throughout B882) for the way it functioned. One of the tutors did not understand why we wanted to call it a name of "Some relative from a country in eastern Europe". It later turned out that they were not aware of who the Wombles were. We realised that this might be a problem with other students as well so we gave UB a FirstClass "Resumé" which explained what the Wombles were and gave my contact details. This information was repeated in a

introductory message posted in the top level conference and in all of the conferences UB had access to.

At this point an important principle became clear. When you are introducing a system such as this you really need to involve those whose approval you need the most. In this case it was not the course team, but the electronic conferencing tutors, who called themselves the 'water rats'. Unfortunately, when the time came to first introduce the idea of the VP to the electronic conferencing tutors they had no idea that this was happening, although it had been planned for almost a year with the full involvement of the course team. The initial reaction was 'no, we don't want it' coupled with 'each year is different, the students don't ask the same questions'. There were several possible reasons for these reactions. First of all B882 conferencing was moving that year from CoSy to FirstClass and this alone was causing a number of problems. Second, they believed that this project was being imposed on them. Finally, there was a certain fear of the unknown coupled with the 'you're going to replace us with computers' fear that AI, or indeed any 'new technology', projects invoke. After convincing them that the project was not being imposed on them and that we did actually need their help, we were then able to discuss the details of the system. We also stressed that the VP could never replace a tutor and was only intended to actively support them; the system would only work in assistance to a tutor and never on its own.

Perhaps the best way to understand the feelings of some of the tutors towards 'Uncle Bulgaria' is to read their own words. I received this message from a tutor who had stopped being a computer tutor before the VP's introduction:[1]

"[Chair] - I know I'm an honorary rat now - but I'll add to the debate! This rigns all sorts of warning bells to me and i for one get extremely irritated by "automatic" messages be they on phone systems or elsewhere. I value Simon's commitment and offer of the programme but on Adult Education grounds there are serious challenges to the concept. Additionally some of the signifcant value of comferencing lies in the fact that another human being, albeit at a distance is taking what i say seriously. To role play for a moment - 'my TMA problems are unique - and an automatic answer is of little vlaue to me - i want to know what a person thinks in reply to what I've said...'

---

[1] All quotes of students or tutors in this thesis are presented as is with no spelling corrections, but with all identifying names omitted.

Well - that should set the ball rolling and If I get an automatic reply - I'll scream!"

This was followed up by another tutor who has since been involved all the way through with teaching involving the VP.

"Well, Uncle Bulgaria is a cat amongst my pigeons! It strikes me as a very exciting idea, fraught with potential, with a rationale that both attracts and repulses me. Reading about it made me feel like an old fogey, too ready with the 'yes, but...'s, while at the same time I can see it could be extremely useful and time-saving, and a fine way of directing attention to the valuable stuff we have archived. I particularly liked the point about it helping students to realise that they are 'not alone' in their problems/situations, and to get ideas about how others had dealt with similar issues before. In that respect, I see Uncle Bulgaria as something like an alumnus or past student.

However, at the moment, like [above], I would have serious reservations about 'him' participating in the main tma conferences. For a start, I do not find answering questions boring (or boaring), and may well choose to answer similar questions differently from how they have been answered (by me or other tutors) in previous years. I do have concerns about the technical/practical reliability of the system, and wouldn't wish the tma conference to become distracted by quirks or technocreaks thrown up by Bulgaria. Because FC is new for some of us (tutors and students ) this year, I'd quite like a run of it without having something else new to be thinking about (like I said, old fogeyism!).

I wouldn't rule out having a *separate* conference (or set of conferences) which Bulgaria could reply in (and there are various ways that could be done), for I'm for experimentation and exploration. However, I'm not keen on incorporating 'him' in the mainstream tma conferences yet - it's hard enough developing sensitive computer-mediated communication as it is, with known, human co-tutors and students."

In the end the tutors expressed four significant fears. The first of these was that it would take over by posting lots of messages to the conferences. The VP only stored nine threads of discussion for TMA 1 and sixteen for TMA 2. This meant that at most it could only post nine messages in response to threads of discussion it identified in the first TMA. It would not post the same message twice. All other messages from the VP had to be explicitly requested by asking it questions. Even if the students were to ask all the questions the maximum contribution during TMA1 would have been 40 messages, and in TMA2 51 messages. As these messages would have been asked for there should be no problem, but even so the total number of messages sent by the VP for TMA2, say, would be less than 10% of the total number messages posted for TMA2 in 1995.

The second fear was that it would put students off. In response it was important to show that the VP was easily identifiable by its name and so students could easily ignore it if they were not interested, and that at most, unless someone was interested enough to ask it a question, the total number of messages it would post was very small.

The third fear was related to one of the VP's design aims; it had been intended to provide an almost real time response in an asynchronous environment. The tutors were worried that it might 'go mad' and start repeat posting to a conference, flooding it with messages. The VP was therefore only ever linked to the conferencing system under supervision so that any erratic behaviour could be caught.

The fourth fear mainly revolved around the effect Uncle Bulgaria would have on the group dynamics of the electronic conferencing. The fear was driven by the feeling that the 'previous topics' which UB identifies will serve to "answer the student's question" or "put them off" so that no discussion of these problems would ever take place. This was definitely not the aim of the Virtual Participant approach. The VP was always intended to be another 'participant' contributing relevant and closely related materials to the discussion. It is possible that some of these negative feelings towards the VP were driven by the tutors' experiences with the previous conferencing system, CoSy, which provided a very flat conferencing structure.[2] With the VP's design, planned for FirstClass, it was far less likely that the small number of VP messages would cause any disruption. As for the content of the messages detracting from the discussion, this was never intended. If anything it was hoped that the VP messages would increase the amount of discussion.

---

[2]With CoSy there was no way to present hierarchical conferences and so distinctions about a message's content had to be made more clearly with the message subject. In addition, the method for reading conferences was very different to that of First Class. Where FC had an all-singing all-dancing GUI allowing you to select and open any messages you wish, and sort the message lists by name, subject, size etc, CoSy provided a different solution. When you entered a conference in CoSy you were given a choice of the available threads. You selected the one you want to read and could choose between seeing all the messages or only those you had not yet read. Once you had read a message you could either make a reply, or go onto the next message in the thread. This reading style made it difficult to skip messages as they would still be presented to you on the screen. It is likely that the tutors still had this interaction model in their mind when thinking about the VP.

Unfortunately, there was no way to find out how a system like this would affect conference-based teaching without actually trying it out. After I had mitigated the worst of their fears, they finally agreed to a trial run, and the VP was approved with manual supervision on the condition that they could request its operation to be ceased at any time.

One final suggestion, which came up on more than one occasion (and which I have always resisted) is the idea of having the VP messages posted to a separate conference. This would defeat the whole object of the system, which was to contextualise its information to the current discussion. By posting the VP's messages to a separate conference the context of the messages would be removed as well as the opportunity for other students to learn from these messages. Also, choosing to notify only those who are actively involved in the thread ignores all those who are just passively reading. Given this reasoning it seems that the current approach is probably the best.

## 4.3 Interacting with the Virtual Participant

Before we dive into the details of what went on during the first year we shall take this opportunity to look at an example of an interaction between students, tutors, and the VP. This example has been taken from the discussions during the first TMA and also serves as an introduction to how First Class appears to the students.

When the student initially logs into FirstClass they are presented with their 'Desktop' showing their personal mailbox and the conferences to which they are subscribed, figure 4.1.

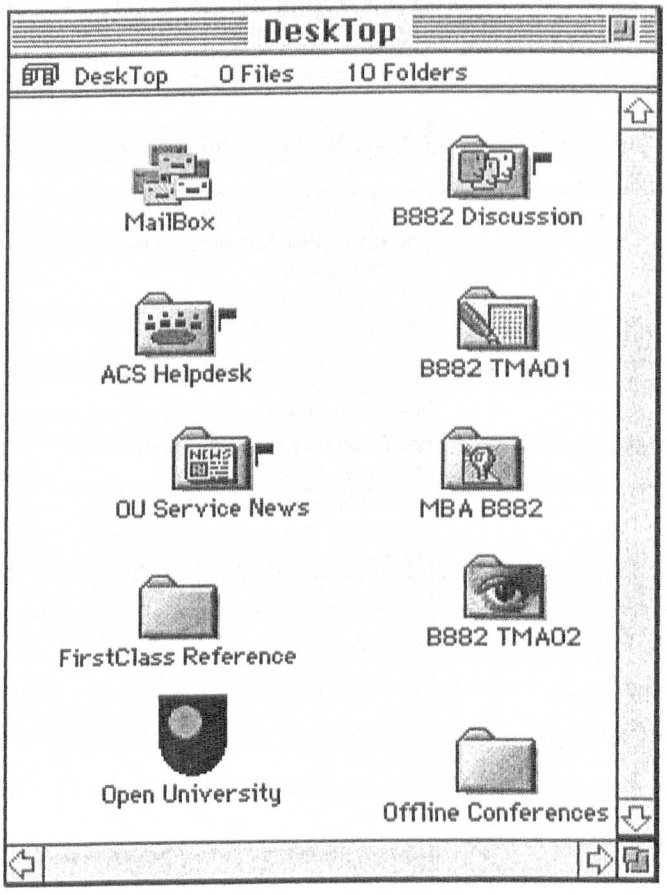


Figure 4.1: A student's desktop in FirstClass. The folders represent conferences. Red flags indicate unread messages.

When a student opens a conference, for example 'B882 TMA01', they are presented with a list of the current messages, most recent first, as shown in figure 4.2. In this view, a number of the messages are greyed out, this means that they are already open. These messages are an example of one interaction by the VP. The following four figures, 4.3 to 4.6, show the open messages from this interaction.

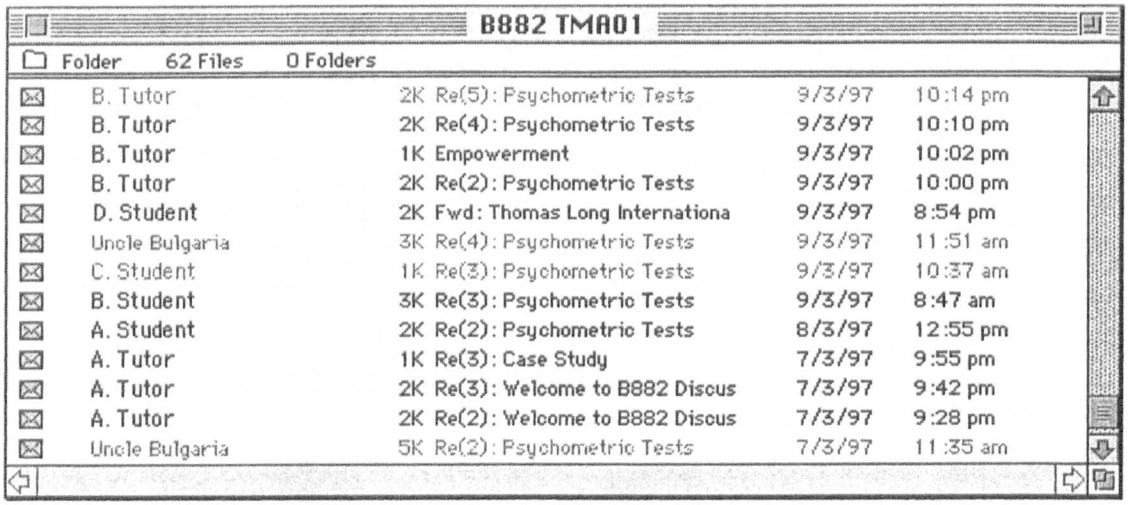


Figure 4.2: The contents of conference 'B882 TMA01'. The names of the contributors has been changed but the distinction between tutors and students has been maintained.

Figure 4.3 shows an overview message from the VP (this was triggered by a student's message not shown in figure 4.2). The messages from the VP are in four sections, shown in figure 4.3. The first section is made up of a 'I think this might be useful' message from the VP and a brief summary. The second section is a selected message from a former student. The third section is a list of questions you might want to ask the VP about this message. The final section is a set of instructions on how to ask the questions in the third section.

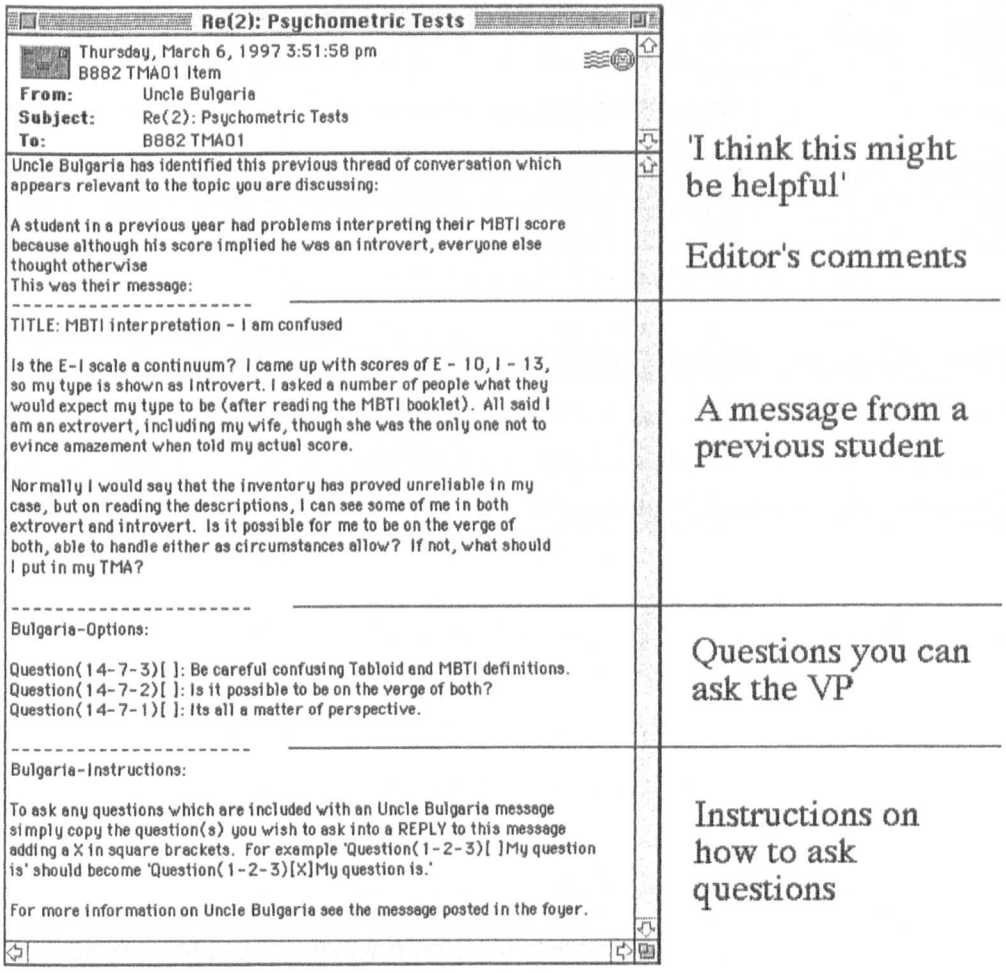


Figure 4.3: This is the first, overview, message from Uncle Bulgaria.

In the second message, figure 4.4, a student has followed the instructions to ask the VP the first question available from the message in figure 4.3. To do this the student first creates a reply to the VP's message; this is automatically addressed to the conference, not direct to the VP. The student then copies and pastes the question they want to ask, placing an X in the box. Then the student can post the message to the conference.

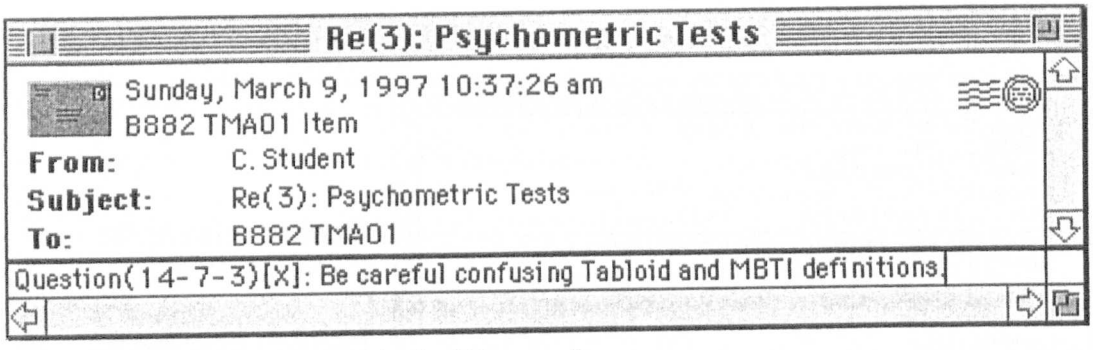


Figure 4.4: A student asks the VP a question.

The VP, reading all messages in this conference, picks up on this question and retrieves and posts a reply to the conference, shown in figure 4.5. In this case the reply contains two messages from previous years, with a short summary for each. There are no further questions that can be asked following on from this message.

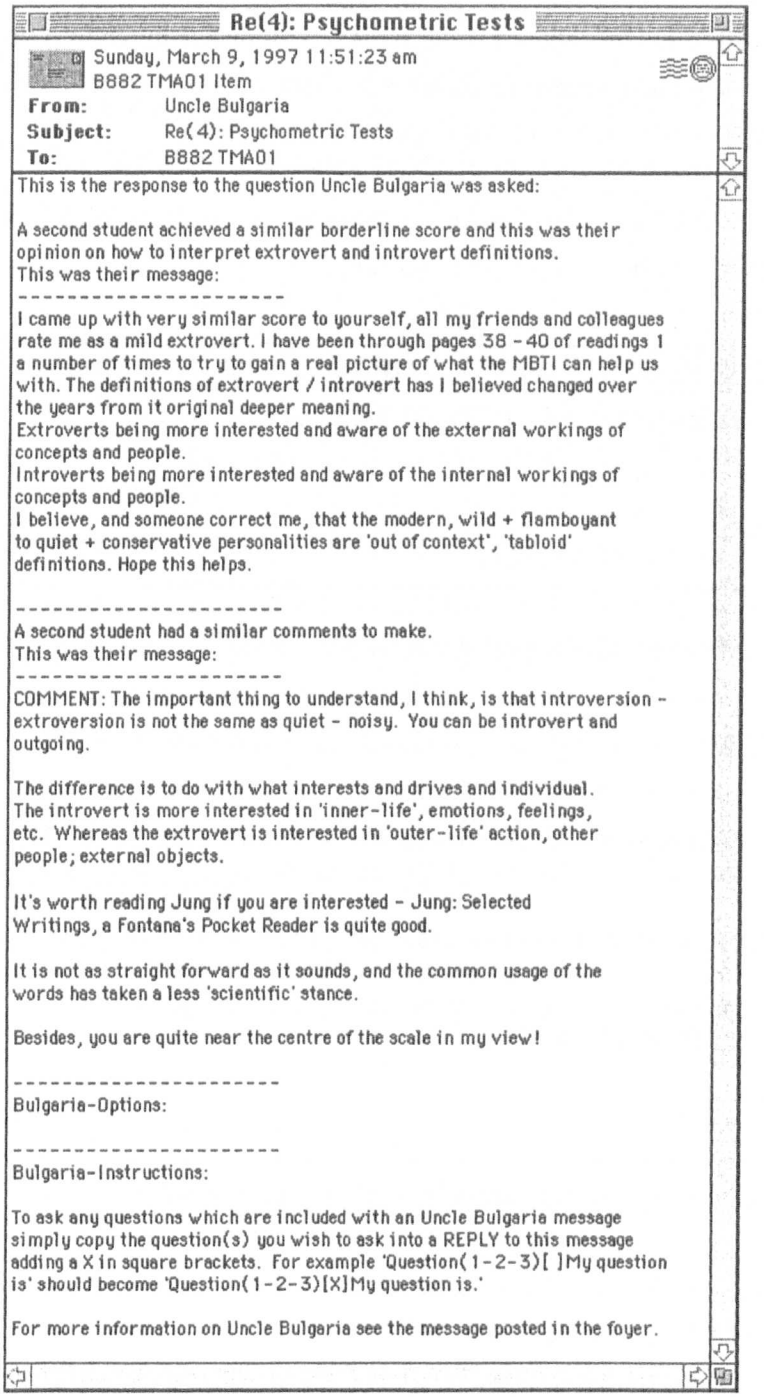


Figure 4.5: The VP's response to the student's question.

In the fourth message, shown in figure 4.6, a tutor has picked up on an important point from the VP's reply shown in figure 4.5. This message from the tutor highlights some important points from the VP's message for the students.

This is a fairly typical VP interaction; in others the students might ask more questions, or more than one student might ask a question.

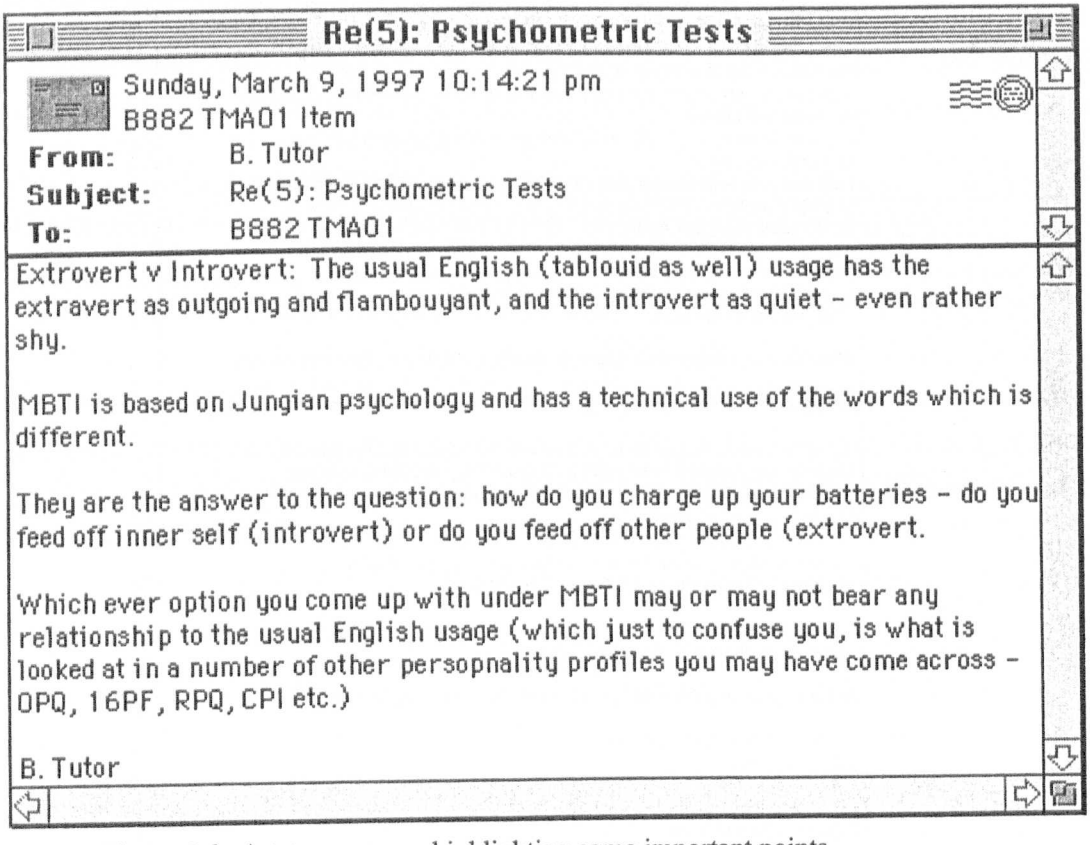


Figure 4.6: A tutor message highlighting some important points.

Note that for technical reasons (imposed by the command line interface), when you view the CLUI you see teletype text, but the standard interface font is Geneva. A PC often presents fonts differently to a Mac, but through the CLUI there was no way to alter the font or size in the first prototype. None of the other normal FC formatting options, such as colours or styles, were available either.

## 4.4 Overview of the use of the first prototype

During the first two assignments the VP made 17 contributions to the discussions. As the experiment progressed some questions received by the VP were via private e-mail; the tutors agreed to let the VP place responses to these in the discussion conference. The breakdown of the VP's contributions is shown in figure 4.7. In total there were 194 messages posted during TMA1 to the assignment and discussion conferences. Uncle Bulgaria posted 12 messages to 4 threads during this time. During TMA2 there were 328 messages posted to the assignment and discussion conferences. Uncle Bulgaria posted 5 messages to 3 threads during this time. The FirstClass history information stored by the VP allows us to study the number of students who have read these messages and thereby may have received some benefit from them. Of the 618 registered users 556 logged in at least once, and of those 363 (65.3%) are known to have read 1 or more VP messages.

We were a bit disappointed in the total number of messages sent by the VP. The retrieval algorithm was weaker than expected, and visual inspection of the conferences revealed some discussions about which the VP had relevant stories which were not retrieved. The same visual inspection confirmed that no cases were retrieved which were irrelevant to the context. Even so, because of this, the VP had a fairly low impact, and this seemed to mitigate the tutors' fears significantly.

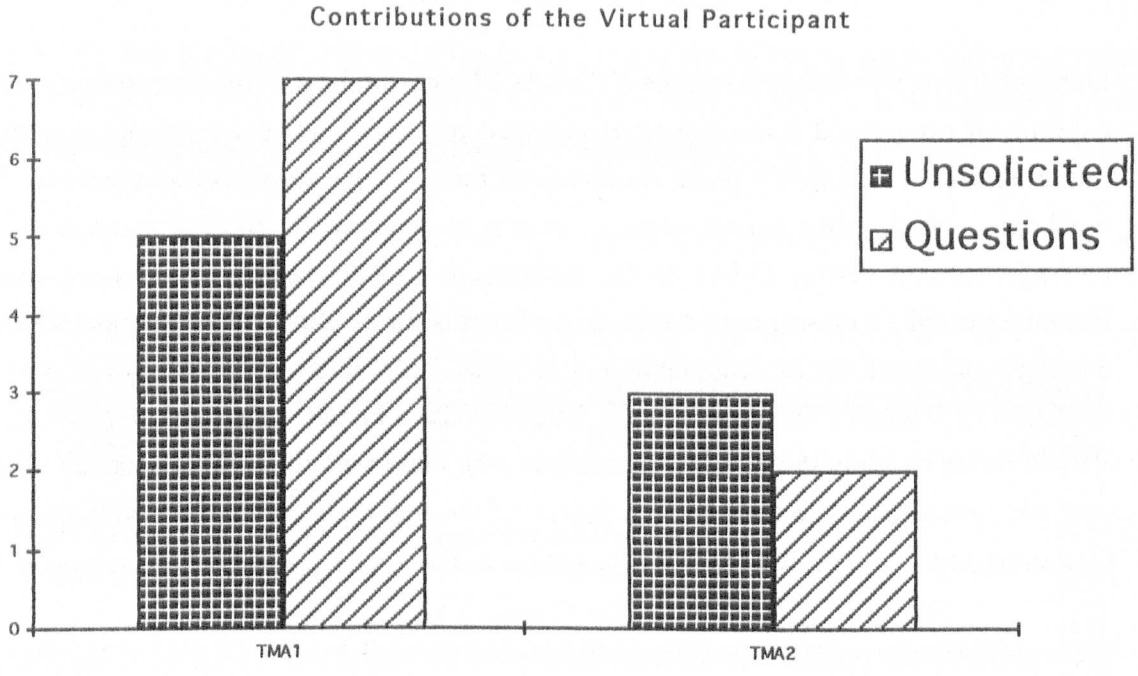


Figure 4.7. The contributions from the first prototype VP in the conferences broken down by TMA and by whether they were unsolicited (U) or questions answered (Q).

## 4.5 The Structure of the discussion threads

The first study in 1997, as seen by the limited number of messages posted and the limited number of interactions experienced, was not particularly large. For this reason, looking for an effect on thread structure is almost impossible due to lack of data. Of the 17 messages contributed by Uncle Bulgaria these were to a total of 7 threads of discussion, 4 in TMA1 and 3 in TMA2. It is interesting to note that the major thread contributed to in TMA2 was called 'TMA02 Networking' - there was a discussion thread with the identical subject label in the second study, and this is one of the main tasks for the second TMA.

| Subject | Unsolicited messages | Questions Available *(for each)* | Questions asked *(for each)* |
|---|---|---|---|
| *In TMA1 UB contributed a total of 12 messages to 4 threads as follows:* | | | |
| "How to do the actual TMA" | 2 | 3, 3 | 3, 1 |
| "KAI tests" | 1 | 2 | 0 |
| "Psychometric tests" | 1 | 3 | 3 |
| "TMA01 Useful Readings" | 1 | 0 | 0 |
| *In TMA2 UB contributed a total of 5 messages to 3 threads as follows:* | | | |
| "TMA02 Networking" | 1 | 3 | 2 |
| "Request for support...." | 1 | 0 | 0 |
| "Part 3&4 of TMA02" | 1 | 0 | 0 |

Table 4.1: The breakdown of UB contributions to TMA1 and TMA2.

In TMA1 there were approximately 36 separate threads of discussion, and the VP posted 12 messages to 4 of these threads. It is impossible to put an exact figure on the number of threads as they can become split by name changes, or by people creating new messages when they meant to reply. In TMA2 there were approximately 85 separate threads, and the VP posted 5 messages to 3 of these threads. This low impact is not necessarily an indicator of poor performance, but rather shows that threshold levels were set deliberately high with the intention of keeping the interactions low and to mitigate the fears of the tutors. Despite this, there was plenty of feedback on how to improve the system so that we could increase the amount of interaction in the second study without fear of causing trouble.

To enable us to look more closely at how interactions with the Virtual Participant progressed we developed a way to represent discussion threads in a diagrammatic form. The diagrams show the logical links between messages (i.e. which message was in reply to another) and the development of the thread over time. The relationships displayed can be thought of as the syntactic relationship between messages, how the contents of the messages refer to other messages would be semantic relationships. Semantic relationships are shown by annotations where appropriate.

The diagrams are developed using the following rules. Each node represents a message, a circle represents a student's message, a square a tutor's message, a rounded square is a

message sent by Uncle Bulgaria and a diamond is a message sent by myself. Each student is represented by a number which is the same across all diagrams on a year by year basis. Each tutor is represented by a letter which is the same on all diagrams across both years. I am represented by the initials SM, Uncle Bulgaria by UB and the Active Archive (the name used by the VP in the second year) by AA. The root node of a thread is placed at the top of the diagram. If a node has more than one daughter, no daughter will follow on from it directly down the page. Time progresses down the page, each node is on a line by itself separated by the same amount from the node on the lines above and below it. When a node has more than one daughter their branches are drawn from left to right in the order in which they are posted. Continuation marks are used as appropriate to keep the whole diagram on one page. Solid lines link messages posted into the conferences, and messages posted privately are represented by dotted links and dotted outlines.

First of all we will look at the thread "Psychometric Tests", shown in figure 4.8, which contains the example interaction presented in figures 4.2 to 4.6. This thread of discussion centred around the two psychometric tests (KAI and MBTI) which the students do as part of TMA1, and what the results of these tests mean. As UB was functioning under supervision it could not run continuously as originally planned, and therefore could not give an immediate response. In this example shown in figure 4.8, UB matches a thread after the second message has been processed and posts its own message to the conference. This message (shown in figure 4.3) is about interpreting MBTI scores and offers three follow-up questions. The discussion then splits into three branches on other aspects of the meanings of the results of the tests. The UB branch continues as shown. What is not shown in the earlier example interaction (figures 4.2 to 4.6) is that later on another student asks the two remaining questions from the message in figure 4.3.

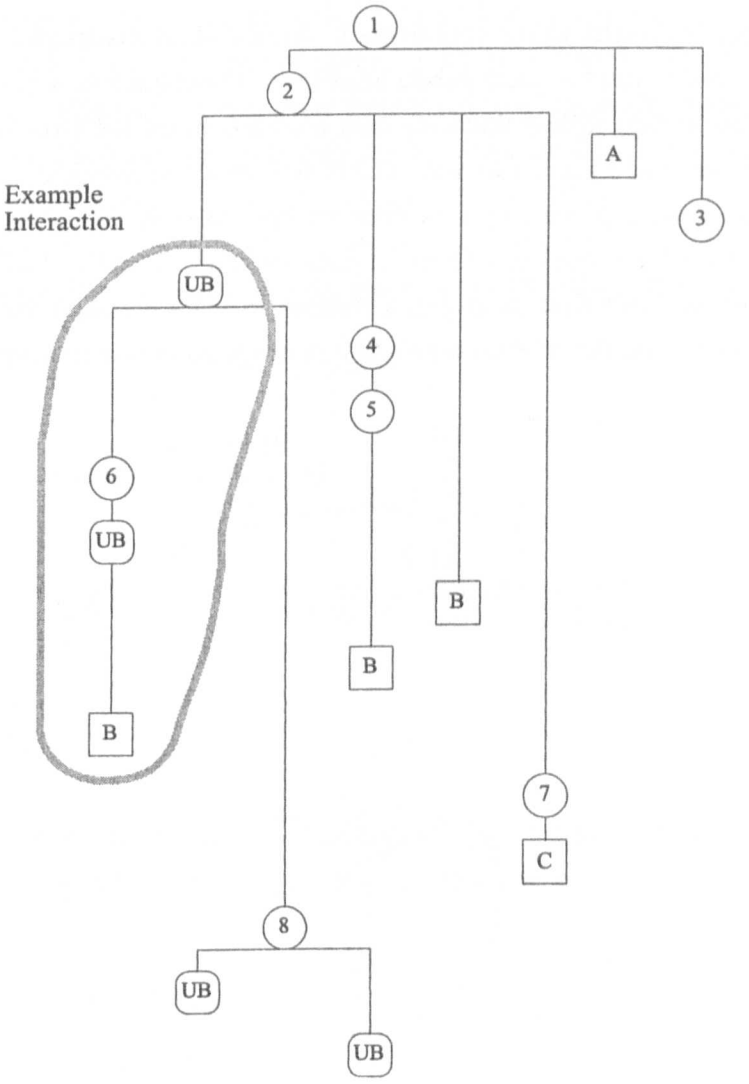


Figure 4.8: Thread diagram - "Psychometric Tests".

For the second thread we will look at figure 4.9, "How to do the actual TMA". This thread was started by a student who was having problems getting good scores on their TMAs for the MBA, compared to those they had got for the Certificate and Diploma courses. UB's message was about how to get started on the TMA and ways of writing. It is of importance to the study because the questions provided by Uncle Bulgaria were of such interest to a number of students that they replied personally, in private rather than in public. UB had not been designed to support this method of interaction (the instructions in each UB and overview message specify this, and anyway the default mode of reply in

FirstClass is directly to the conference). After a short discussion with the tutors it was decided that I could request public answers to these questions myself. We had a tutor (not a 'water rat') follow these up with a comment on the structure/presentation, then another tutor and I replied to that. Uncle Bulgaria also matched another message which was a tip from a previous year's tutor on how to start on TMA1. This message was followed up by a question from another student. Finally, following up the tutor's comments on UB's first messages, a student posted a message commenting on some of the advice in both the original tutor's first message and those received from UB.

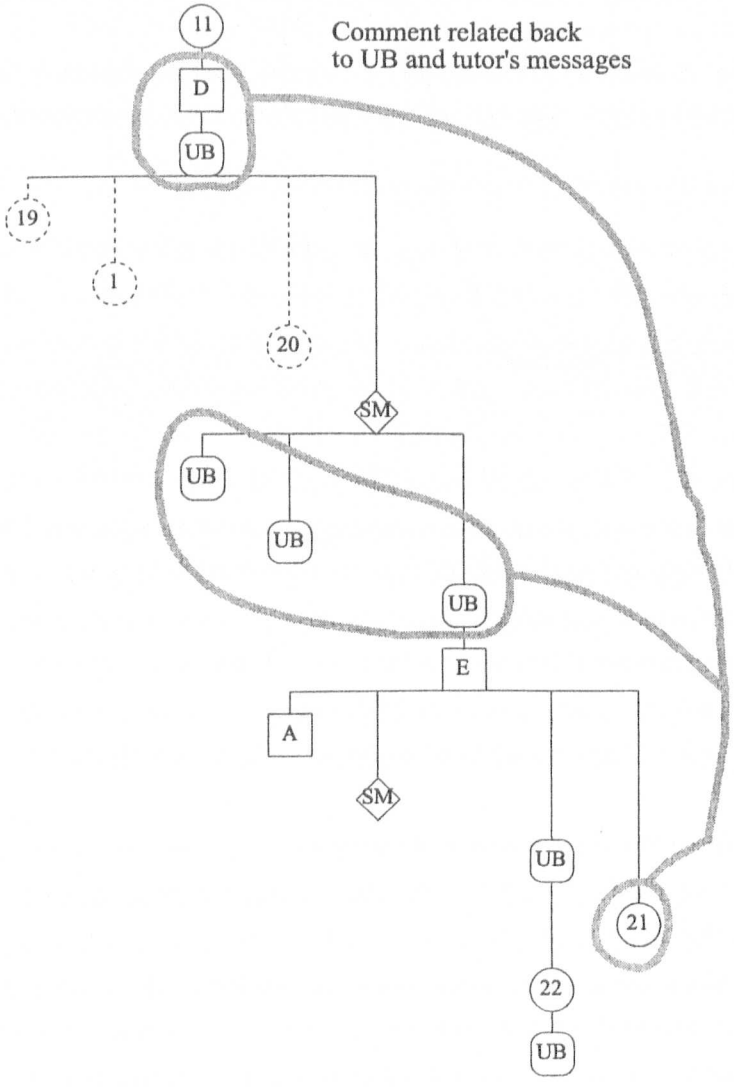


Figure 4.9: Thread diagram - "How to do the actual TMA".

The last two threads in TMA1, figure 4.10, "TMA01 Useful Readings" and, figure 4.11, "KAI tests", are included here for completeness. First we will look at the thread "TMA01 Useful Readings". This was a fairly short thread and was created by a student who was trying to get their TMA started by listing the readings for Block 1 which were most relevant. A couple of other students followed up with other suggestions and UB posted a slightly off topic message about a previous student and their first ideas for TMA 1. This was what the student was working towards, but it was not related to the readings.

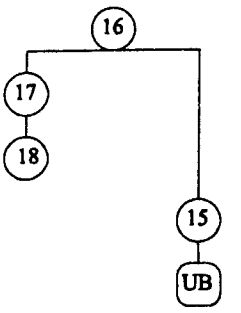


Figure 4.10: Thread diagram - "TMA01 Useful Readings".

The last thread, shown in figure 4.11, "KAI tests" was started by a student when they received their KAI results. Most of the discussion revolves around what scores each student got. This then progresses onto speculation that B882 attracts high KAI scores and the tutors addressing the usual misconception that a low KAI score means low creativity, rather than its true meaning as a measure of how people are creative. Uncle Bulgaria's message to this thread was from a previous year's student speculating on their KAI score and mentioning that they had met a number of high KAI scoring managers on B882 and they all appear to be looking for new jobs, because their high KAI score means that they cannot understand, or be understood by, their colleagues.

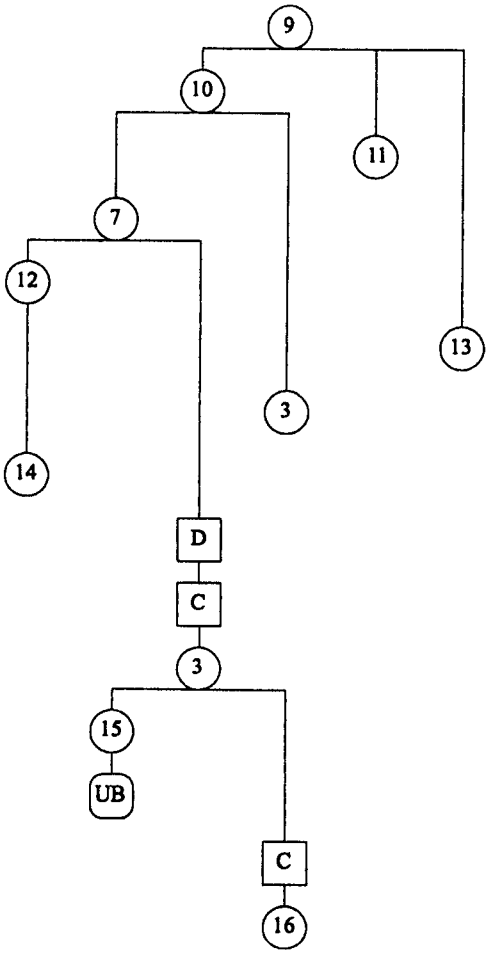


Figure 4.11: Thread diagram - "KAI tests".

During the second TMA Uncle Bulgaria was much less active. There were several reasons for this, the main one being that the thresholds were deliberately set very high, but there was also a large increase in the use of the newly released offline reader making it harder to track people following threads. The three threads matched by UB show a weaker performance than in the first block. This was almost exclusively due to the thresholds being set too high for the second block. A number of threads were matched near to the threshold, but very few actually exceeded it. In the thread "Part 3&4 of TMA02" one of the students was discussing their plans for a brainstorming session at work to solve a problem her department was having. The thread matched by UB was on the Delphi Method and a student's previous experience with it. This was not quite on track, although was related to their discussion of techniques they could use. The

"Request for support...." discussion revolved around a student with the problem that they were not able to get a group of others at work involved in a brainstorming session. The thread matched to this was a one-off 'light humour' message posted by a student in a previous year including a flow-chart described as 'the ultimate problem solver'. This was a completely irrelevant message, but something I left in the case base to see what, if any, reaction was provoked by it. No comment was made by tutor or student.

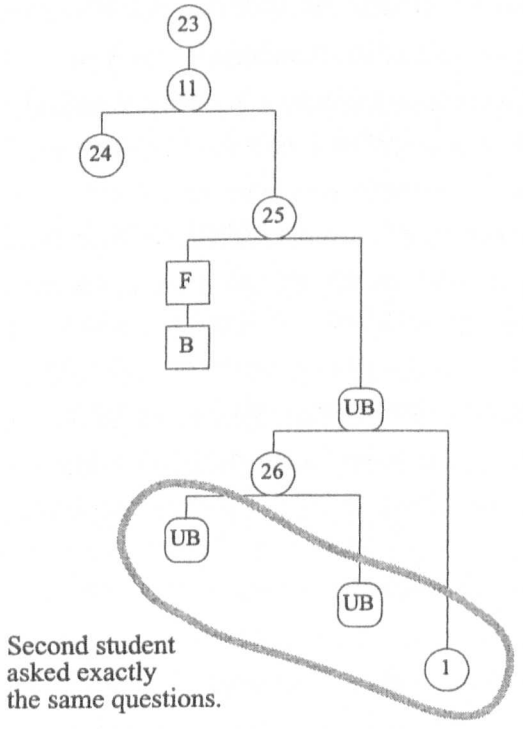


Figure 4.12: Thread diagram - "TMA02 Networking".

The thread "TMA02 Networking", shown in figure 4.12, was a discussion by a number of students working on the fourth part of the TMA trying to find ways of using others in your problem solving tasks. A thread with exactly the same name, and very similar content, also occurred in the second year study, supporting the hypothesis that the same issues do appear in different years. In this discussion the students were looking for others to comment on their ideas, approaches and techniques. Uncle Bulgaria offers a number of cases giving help on the topic of brainstorming, particularly for busy people or those away from the office, reflecting on their discussion on how to manage a brainstorming session. One student asked two follow up messages to UB, and after these were posted another student later asked exactly the same questions.

## 4.6 Summary

In this chapter we have seen how the VP was introduced to the tutors and students. We have looked at the behaviour of the system in a real world situation and how the tutors and students interacted with it. The first prototype was not perfect (indeed it was never expected to be) but there was a lot that could be learned by observing the students' and tutors' reactions to it. It was obvious that there needed to be improvements in the presentation and a rethink of the interaction. An important observation was the hostility that the name could produce amongst a minority, especially a name with attached cultural baggage. Overall, this study was rather quiet, with very few messages posted by the Virtual Participant, but there was enough to make an assessment of what needed improving. The matching of current discussions to past threads was successful enough to show that messages were mostly (if not always) on target. This seems to justify the approach taken, demonstrating that it is viable. Another success was that the students did actually ask questions of the system, although there did seem to be many who preferred to ask questions privately. Pulling together these observations and the questionnaire feedback discussed in the next chapter helped to improve the prototype of the Virtual Participant for the second study, which I will discuss in chapters 6 and 7.

# Chapter 5: What the students thought

## 5.0 The first prototype, what did they think?

This chapter discusses the initial study into the Virtual Participant. First, we look at the survey we conducted and how and why we selected the students that we did. We then draw out the important points from this survey, summarise the comments made by tutors and students over the course of the first study, and look at the conclusions we can draw from it. Finally we look back over what has happened and what we have learned, and then look forward to what we would like the second prototype to achieve.

## 5.1 Survey about the first prototype

As this research has been intimately involved with an existing course it has always been necessary to balance the needs of the course and those running it against the needs of the research. With this balance in mind the survey contained three sections; it can be found in appendix 3. The first section was about First Class and what the students thought about it, the second was about the course-related web site, and the third section about Uncle Bulgaria. The inclusion of the first and second sections were mainly to provide information for the course team and the 'water rats' and to help justify the survey. These two sections also served to provide additional background information on the students' opinions. Over and above the survey, there had already been some feedback from the students and tutors in the conferences, and more could be deduced by observing the interactions. Because of these other sources of information it seemed appropriate for this particular survey to concentrate on heavy users of Uncle Bulgaria, who were often heavy users of First Class.

The main rationale behind the questions relating to Uncle Bulgaria was to look for evidence regarding some of the more important issues raised by the Virtual Participant. The key things I wanted to know from the students were:

- The student's opinion of the name (Uncle Bulgaria)?
- Did it put them off from reading messages or contributing to the conference?
- Was the content of the messages useful?
- Should it continued to be used?

I also wanted to know answers to more implementation-driven questions:

- Would they like access to the whole knowledge base?
- Would they prefer to ask it questions privately?
- Were the answers really proving relevant?

There were no other high level goals involved in the use of the questionnaire, as (from the other feedback I had received and observations I had made) I already had a fairly clear idea of the path the development would be likely to take. Instead I was rather looking for support for other - often anecdotal - evidence collected during the study.

Each section of the questionnaire asked the students to rate a number of statements against the scale strongly agree, agree, disagree, strongly disagree. The end of each section provided a space where students could add any other comments they wished. The format chosen was to offer the students a number of statements on each topic and ask them for their opinions. The design and format of the survey was influenced by the course team.

We surveyed 104 students from the group of 363 who read at least one VP message. Students were selected if they had read 9 or more VP messages, making a sample of 120. We were required to eliminate 16 of these (leaving 104) because they had already been surveyed by the university this year. In effect, then, we only surveyed the heaviest users of the VP.

The results of the section on the web site have no relevance to the thesis work and so are omitted. The statements on FirstClass were as follows:

1. It is easy to identify the right conference
2. It is easy to find out what each conference is about
3. The content of the facilitated discussions is valuable
4. The content of the free discussion areas is valuable
5. Support by the special tutors ("water-rats") is valuable
6. Overall, I was very satisfied with B882 conferencing

7. I found the e-mail facility very useful
8. I found the wider MBA conferences very useful

The full results can be seen in figure 5.1, and the summarised results in figure 5.2.

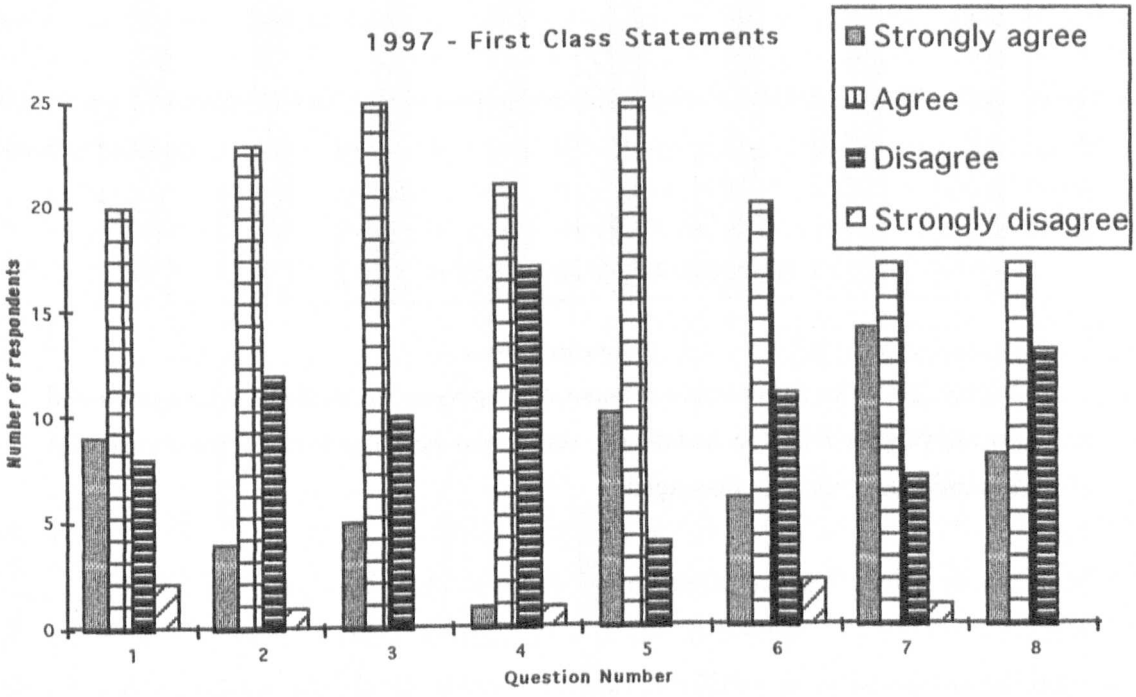


Figure 5.1: The number of responses for each statement about First Class, note that not every student answered every question.

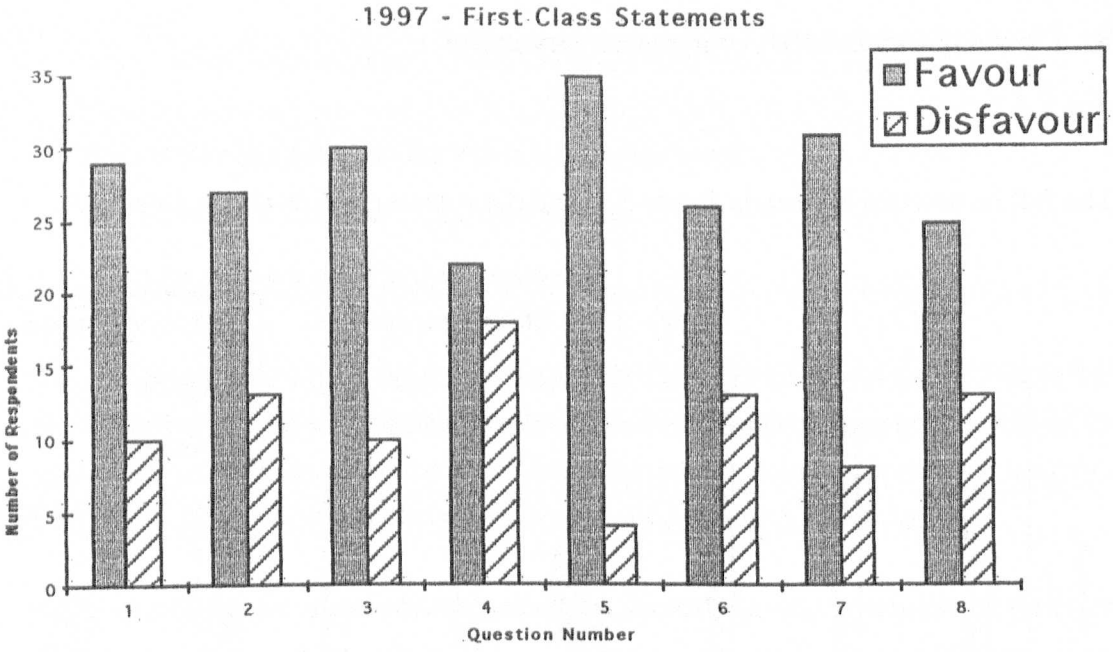


Figure 5.2: The summarised number of responses for all those who agreed and strongly agreed to the statement, versus the combined value for those who disagreed and strongly disagreed.

These results show that the students are positive about the majority of the statements. Question 4, about the content of the free discussion areas has often proved a point of contention and usually, as seen here, there is about a 50/50 split between those that find it useful and those that don't. In responding to question 6 two thirds of students said they were satisfied with B882 conferencing overall, obviously there are plenty of opportunities to improve on this.

The section on Uncle Bulgaria presented the following statements:

Uncle Bulgaria's messages...

1. contained useful information
2. were relevant to the discussion
3. answered questions I wanted to ask
4. provided a different viewpoint
5. put me off sending my own messages

6. I would prefer to ask UB questions directly by private e-mail, rather than through the conference
7. The name Uncle Bulgaria is a good choice
8. Uncle Bulgaria should continue to be used on this course
9. I would like access to all the information UB has

With these statements, with the exception of statement 5, we were hoping for agreement from the students.

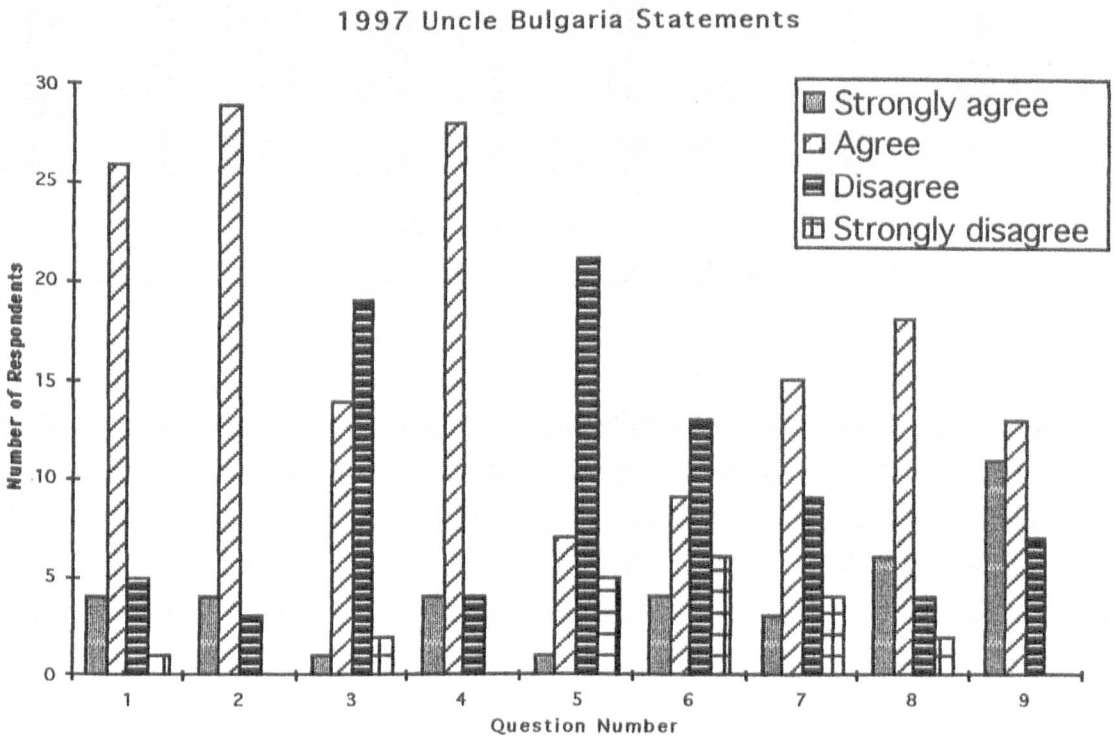


Figure 5.3: This chart shows the number of responses for the statements about Uncle Bulgaria.

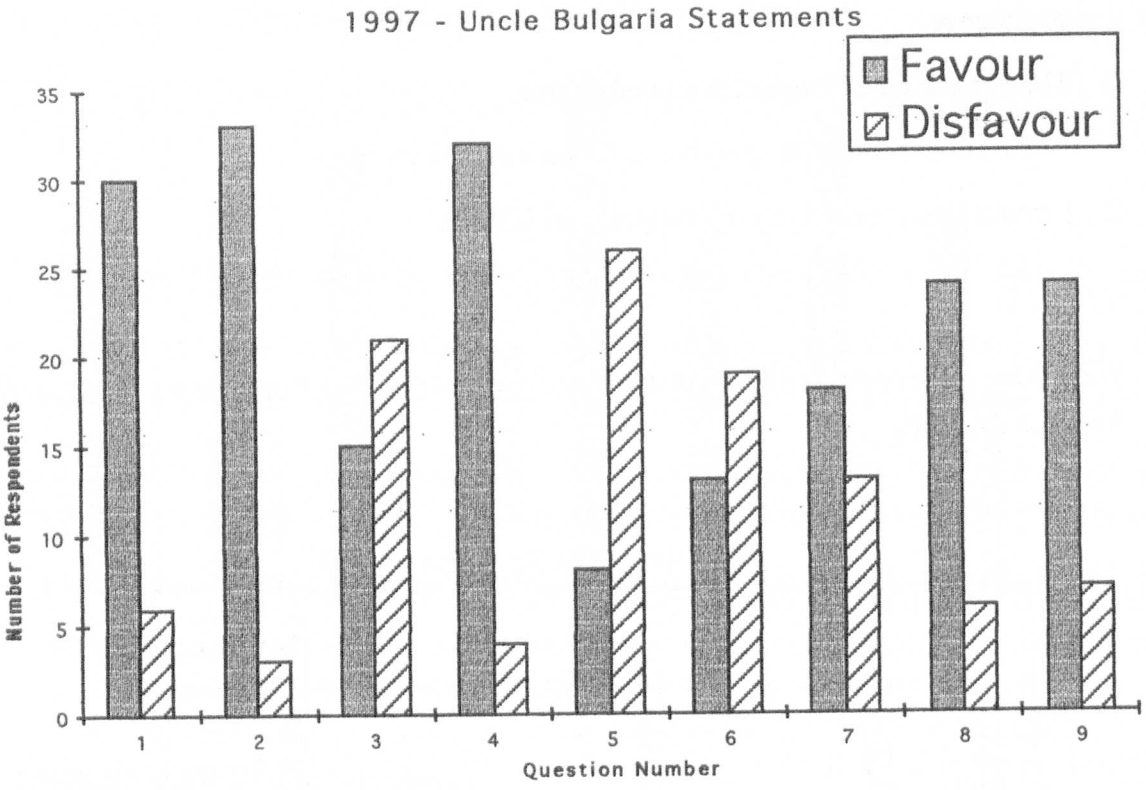


Figure 5.4: This chart shows the agree/disagree summaries for the statements about Uncle Bulgaria.

When we look at the results we see large positive agreement with statements 1, 2, 4, 8, and 9. With question 5 we saw a large disagreement, as we had hoped. The outcome of conducting 1 by 2 chi-square tests on each of these results can be seen in table 5.1. Apart from statements 3, 6, and 7 there is a high significance for each result. The full results are shown in figures 5.3 and 5.4 taking the same format as 5.1 and 5.2 respectively.

| | Chi-square | Significance |
|---|---|---|
| Uncle Bulgaria's messages... | | |
| 1. contained useful information | 16 | p<0.001 |
| 2. were relevant to the discussion | 25 | p<0.001 |
| 3. answered questions I wanted to ask | 1 | p>0.1 |
| 4. provided a different viewpoint | 21.8 | p<0.001 |
| 5. put me off sending my own messages | 9.53 | p<0.01 |
| 6. I would prefer to ask UB questions directly by private e-mail, rather than through the conference. | 1.12 | p>0.1 |
| 7. The name Uncle Bulgaria is a good choice | 0.80 | p>0.1 |
| 8. Uncle Bulgaria should continue to be used on this course | 10.80 | p<0.01 |
| 9. I would like access to all the information UB has | 9.32 | p<0.01 |

Table 5.1: Results of 1 by 2 chi-square tests on responses to Uncle Bulgaria statements.

When we look more closely at statements 3, 6 and 7, figures 5.5 and 5.6, a clearer picture emerges. In the case of statement 3 ("UB answered questions I wanted to ask") there is a pretty even split between agreement and disagreement. It is likely that some respondents may have been slightly confused by this statement taking it to mean 'answered questions that I asked' rather than 'answered questions that I wanted to ask'. Given the strong agreement with statements 1 and 2 we feel that the response to this statement is not a bad reflection on the VP.

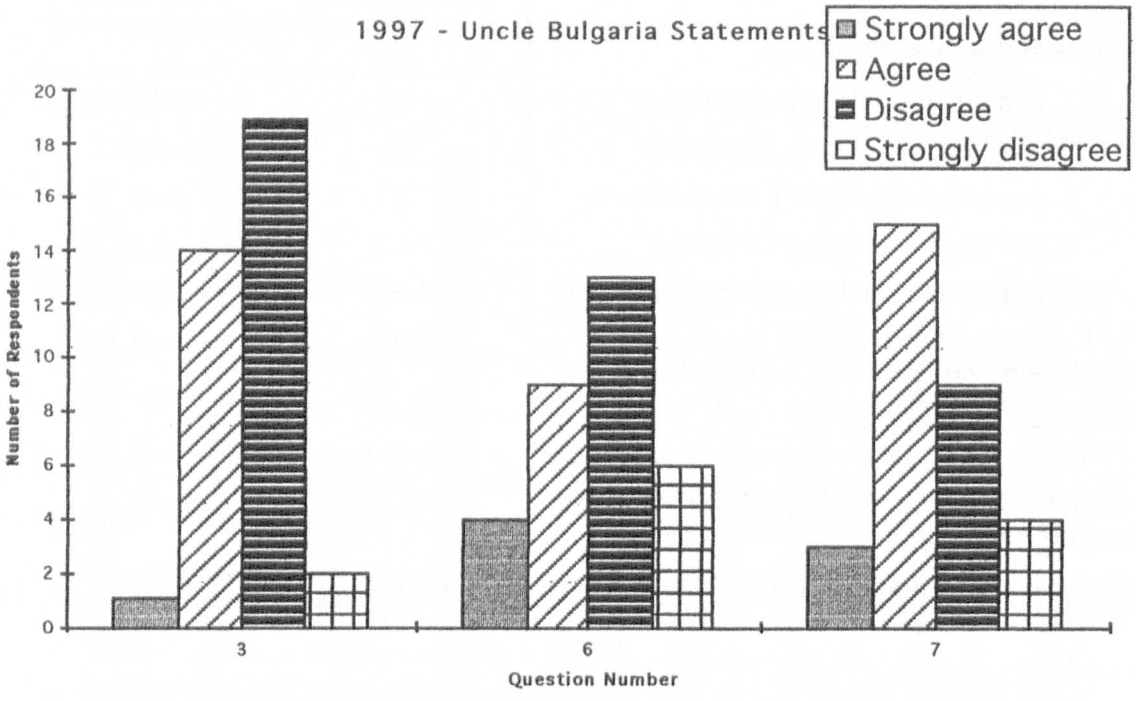


Figure 5.5: This chart shows in more detail the specific breakdown of results for questions 3, 6, 7.

Statement 6 ("I would prefer to ask UB questions directly by private e-mail, rather than through the conference") is a different matter. In this instance we see a more 'normal' distribution with more students agreeing or disagreeing strongly. One comment received on the questionnaires ([10] in appendix 5) indicated indifference to the approach, but this was from a participative student who never asked UB a question! From the responses to this statement and the other feedback I received it seemed that allowing students to ask questions directly (privately) is the way to go, as only allowing questions to be asked in the conference is often a distraction and generates unnecessary excess messages. It is also quite likely that this is a reflection on the natural reticence of students, unwilling to be seen asking questions of a machine.

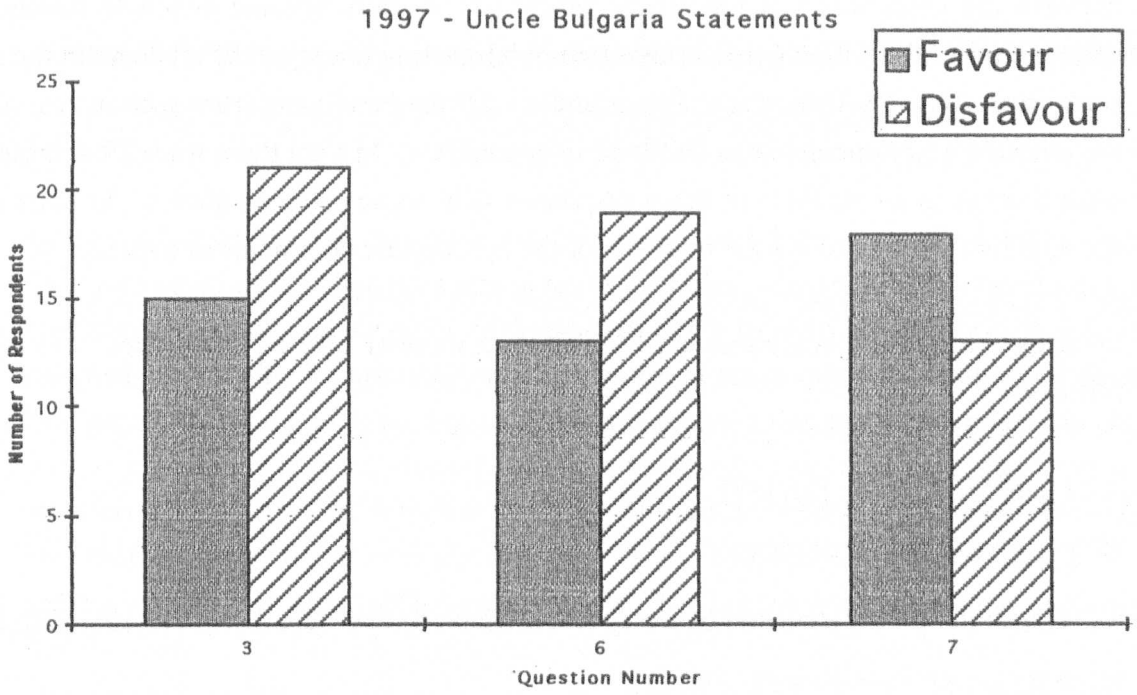


Figure 5.6: This chart shows the agree/disagree summaries for statements 3, 6, & 7 about Uncle Bulgaria.

Statement 7 asks for an explicit opinion on the appropriateness of the name of the VP. The majority of the opinions were favourable or indifferent; "*No feelings either way its just a name to recognise. Made me think of parrots!*" (typical questionnaire comment). However there was a clear negative reaction from some users and this is discussed later in section 5.4.

## 5.2 First Class comments in the questionnaire

In each of the three sections of the questionnaire we also asked for general feedback comments from the students. The second section, on the web site, is irrelevant to this work, and was only intended for use by the course team. The feedback on the students' opinions of First Class and their opinions of Uncle Bulgaria are both relevant to this research. Here we will discuss the feedback comments we received from the students. In this section we will start with the comments relating to First Class. In the next section we go on to look at the specific comments on Uncle Bulgaria.

In the first study all the students had previously used CoSy, whereas in the second study they had used First Class and had received training in it. Another clear difference

between the years was that the offline reader for FC only became available during the second study. With this in mind there was much feeling amongst the students in the first study year about the lack of an offline reader. All the comments from section two of all the returned questionnaires can be found in appendix 4. In total there were 29 comments, what follows is an analysis of these responses with some selected quotes. In each case the numbers referred to are the numbers of the appropriate comments in appendix 4.

Looking through the comments on First Class a number of clear categories emerge, as follows.

**A.** I hate First Class: [2], [17]

**B.** I like First Class: [1], [23]

**C.** The conferencing structure is messy and difficult to use: [5], [7], [10], [13], [22], [24]

**D.** Who are the 'water-rats'?: [3], [7]

**E.** Poor content, discussions off topic: [4], [7], [10], [11], [14], [18], [20], [29]

**F.** Positive about computer-mediated communication: [1], [4], [6], [15], [21], [23], [25]

The remaining comments: [8], [9], [12], [16], [19], [26], [27], [28], are hard to classify but are generally about lack of time to participate in conferencing, or problems with the technology.

Categories A and B show how the change from CoSy to First Class has affected some students. Those who hate First Class do so because there is no offline reader, so there is an increase in telephone bills, but they also found it more difficult to use, as there is no ability to thread discussions. Those who liked First Class did so because they were "... *suprised to find how useful & easy to use it was*" [23] in comparison to CoSy. Because the GUI interface FirstClass provides has always been preferred by students over the vt100 style of CoSy there has been an overall increase in the numbers of students using conferencing since the systems were switched.

Category C has six comments on how difficult it was to find the correct conference. Some of the comments indicate this was often due to an unfamiliarity with First Class: "*Finding a conference is OK once you are familiar with the system but people non computer literate find difficulty. some important course topics are buried as sub-conferences and are thus less obvious. once found you can use the alias to put onto your*

*desktop."* [7] There was also some criticism about some of the conferencing being split up on a regional basis: *"Surely the big advantage of cmc is that location does not matter."* [13].

Category D only received two comments, but it is important because of what it reveals. The 'water-rats' is the colloquial name given to the tutors' who are based in the electronic conferences. A lack of recognition of these tutors by the students (especially as they were sampled specifically because they were heavy users) raises questions about their input into the conferencing process. For the Virtual Participant this can be interpreted in two ways. The first is that the students will pay no attention to it, just like they pay no attention to the tutors. The second is that they will consider what it has to say in the same way they consider what everybody else says. These two interpretations are really like arguing whether a glass is half full or half empty. The key point to take away from this is that the students are evaluating what they read and if they are treating the tutors like any other participant then it also seems probable that they will treat the VP in the same manner.

Category E received the most comments with eight. These were all criticisms that there was too much chat in the conferences, too much off topic discussion, and not enough management of the discussions. One student states *"There tends to be a lot of trivia - unfocused dialogue by users in conferences intended to be on course topics. chat areas are the correct place."* [7] These criticisms are of most importance when you realise that the overall effect is to reduce the motivation of the students: *"Disappointed by lack of substance in conferences & poor response. Felt I put in more than I got out of it so my interest waned later in the year."* [14] There is a natural drop off during TMA 4 but during this study year the drop during TMA 3 was greater than expected. This lack of motivation and feeling of frustration with the conferencing can be seen in the comments in both this category and category C.

Finally, in category F, the comments about how important CMC is were another positive sign. Conferencing of this kind can be very helpful for distance learning students in promoting a stronger feeling of community: *"It was useful to keep in touch with other students, and also to feel that support was always available."* [6]. And for some students it is a vital factor: *"The reason I chose the OU is it is a distance course which I can do to fit in around a full time demanding job and a small demanding son! to suit me best I need to be able to study when I can - any time day or night. The conferencing system is*

*wonderful so that - I am not a good attender of seminars and tutorials because they eat into my time with my son, so the conference provides me with access to other students and ideas. So I am sold on it."* [25].

Overall the students were very positive about the ideas behind conferencing, but there were many problems which put them off in practice. Not realising (because there is no obvious distinction between any users) who the computer tutors actually are is a major problem for the use of this medium for teaching. Other problems with messy and confusing conference structure and badly managed discussions are for the tutors and the course team to deal with. The main reason for a lot of the student problems, I believe, was their (and the tutors) inexperience and lack of training with First Class. This was not the case for the second study.

## 5.3 Uncle Bulgaria comments in the questionnaire

In this section we review the questionnaire feedback comments received about Uncle Bulgaria. These comments add to the other comments we have received in the conferences and informally from the tutors as the trial progressed. We look here at the groups into which these comments fall and in the next section we pull all the feedback and study evidence together to look more closely at how the system should be revised.

Included in the following analysis are the comments of one tutor on the style of the messages from UB, even though these were made directly in the discussion conference. Additionally, one of the tutors started a discussion about what the students thought of UB. This took place in a conference outside those which UB had access to and therefore could not be captured. The informal feedback to me from the tutor was that the discussion was limited to one student complaining about the name and not knowing what a Womble was and who they quoted as saying the messages from Uncle Bulgaria implied *"you silly people, this topic was discussed and sorted out ages ago"* (in appendix 5 as comment [14]). They quoted another student as observing that *"Some of the contributions are appropriate, others are a tad on the side, but this may be useful as it may be providing a view from another angle"* (in appendix 5 as comment [15]). Two or three other students joined the discussion reminiscing about watching the series with their children; they also observed that the messages 'didn't look very nice'. The fact that there were generally very few comments about it directly in the conferences I take to be a good sign, but it does make it very important to pick up on what few comments there were.

The full comments from section three of all the returned questionnaires can be found in appendix 5, along with the message from the tutor on presentation (comment number [16]). In total there were 13 comments from the questionnaire, two extra ones from students, and one extra one from a tutor. What follows is an analysis of these responses with some selected quotes. In each case the numbers referred to are the numbers of the appropriate comments in appendix 5.

Like the FirstClass comments above we can classify the feedback on Uncle Bulgaria into a number of categories:

**A.** Positive about UB: [2], [6], [7], [8]

**B.** Unclear about the UB system: [4], [5]

**C.** Comments on the naming: [1], [3], [6], [11]

**D.** Presentation unclear: [5], [7], [16]

The comments in category A showed that there were students who thought that it was a good idea, for example: *"I think UB is and interesting experiment and what better course than B882 to experiment on?"* [6] and another student who felt that *"...UB helped structure/ensure good quality information in the conference."* [8] So the idea and the content were seen positively by some students. The other three categories point to areas for improvement.

In category B we find comments from students who are *"... a bit confused what it was all about"* [5], *"... , and did not fully understand the concept - where was it explained?"* [4]. There were messages posted explaining what UB was about at the very top level of conferencing, in the coursework conferences, and in each of the conferences in which it operated. These comments indicate that this was not enough and more effort should be made to raise student awareness of the system. They also, perhaps, reflect some of the problems students were experiencing with the new conferencing system.

Category C comments were on the name we gave the system. These comments ranged from: *"I haven't a clue why Uncle Bulgaria is called Uncle Bulgaria!"* [1] (although this was detailed in its résumé) to: *"I feel that the name had an adverse affect on what the objectives of UB were"* [3]. Another student commented that: *"It needs an unusual name so it is clear that it is not a student or a tutor"* [6] (it appears that 'Uncle Bulgaria' is not

unusual enough). It is clear from these comments, and the responses to the statement about the name in the questionnaire, that something needs to be done to clearly distinguish the system from a normal user, by its name. The selection of a good name will also serve to reduce any negative effects on the system's objectives.

Category D comments refer to the presentation of the content in UB's messages. This was also commented upon within the FC conferences by a tutor *"...Uncle B. is not very structured. Could he have some paragraph numbers or something?"* [16] in direct response to one of UB's messages. The cause of this problem was the limitations placed on the system's interactions with FC because of the command line interface. This imposed a fixed font size and style, and also required each line to be properly terminated. Because of this, viewing UB's messages through the graphical interface made them look slightly odd, and certainly rather uniform. In comparison to this all GUI users could (and did) freely use a mix of colours, fonts, and styles.

These category groupings are quite simplistic and we received only 13 feedback comments. However it does seem likely that the points raised are indicative of general problems with the system, and are points that should be addressed in a revised design. The next section pulls together the comments reviewed in this section and our own observations during the first study. We then go on to specify the key points that we have learned from this experience and how these and other ideas can feed into the second prototype.

## 5.4 Drawing together the feedback received for Uncle Bulgaria

In addition to the questionnaire results we have student comments and some feedback from the conference discussions. Combining results from all the feedback received there are four clear areas which need to be addressed: The name of the system, the presentation of the messages, the content and context in which they were presented, and finally occasional confusion of the users.

*Name*

Some users did not like the name 'Uncle Bulgaria'. They felt that messages from that user implied *"you silly people, this topic was discussed and sorted out ages ago"* [appendix 5, 14] and that *"the name had an adverse affect on what the objectives of UB were"*

[appendix 5, 3]. One student pointed out that *"it needs an unusual name so it is clear that it is not a student or tutor."* [appendix 5, 6]. Given the feedback we received and the specific dislike expressed we believe this is due, not to the content of the messages, but to the 'avuncular' presentation of the name. The results of the survey, statement seven in figures 5.5 and 5.6, show a reasonably even distribution in opinion, but it is the opinion of those students who really disliked the name that we must address.

*Presentation*

As mentioned before, the communication protocol between the VP and FirstClass was limited. The students' interface to the conferencing system provides much richer formatting. People like to skim messages and plain text does not easily emphasise important information. When the first UB messages were posted to the conference tutor E, who is not a computer tutor, replied to the whole conference commenting that the format and structure of the messages could be improved. One student commented: *"useful at times: continue with it, more 'white space' on sum please"* [appendix 5, 7]. Another deficiency in FirstClass is that it does not support threading so when more than one question was asked all replies would appear grouped together. Add this effect to the problems some students had in understanding the messages from UB and you find a student who *"was a bit confused what it was all about and had difficulty following the thread of discussion"* [appendix 5, 5].

*Content and Context*

With the VP we were aiming to provide *content* which is useful to the students in the current *context*. In this way we can enhance the learning for the students. With this section we want to discover what the students actually thought about these factors. One B882 student said that *"Some of the contributions are appropriate, others are a tad on the side, but this may be useful as it may be providing a view from another angle"* [appendix 5, 15]. So certainly for them the context was not always appropriate, but they did seem to find the content useful. Another perspective from a surveyed student was that *"I felt UB helped structure/ensure good quality information in the conference. One of my criticisms of conferencing is that tutors (as they would in a real tutor group) do not control poor quality student input"* [appendix 5, 8]. For this student the VP had the effect of actually *increasing* the quality of information available to them. Another student felt that the messages from UB *"...tended to reflect points raised by other contributors to ask*

*questions to stimulate debate rather then be informative"* [appendix 5, 9]. A goal of the VP is not to answer questions, but enable reflection on them. While this may feel frustrating for more goal-directed students it is hoped that they will learn more this way.

*User Confusion*

There are some clear indications that a number of students did not know what Uncle Bulgaria was, or what its intentions were. Some aspects of this problem could be overcome by better training in the use of FirstClass, for instance knowing how to access a résumé. Other aspects may be addressed by giving it *"an unusual name"* [appendix 5, 6]. We received a couple of comments in our questionnaire about this: *"Did not find of any great use, and did not fully understand the concept - where was it explained?"* [appendix 5, 4], and *"Was a bit confused what it was all about and had difficulty following the thread of discussion"* [appendix 5, 5]. These issues are well summarised by one questionnaire comment: *"It needs an unusual name so it is clear that it is not a student or a tutor. Unfortunately there are always some who do not read the messages explaining what UB is and then post a message asking 'who is Uncle Bulgaria?'"* [appendix 5, 6]. This has encouraged us to make the role and behaviour of the system much more clearly defined.

Finally, and very importantly, we also received some feedback from the tutors who felt that they would like to know more about what information the VP had, and its current state of operation. We feel that the lack of this information may have led to a certain 'fear of the unknown' from the tutors.

*Summary*

Our introductory work has raised a number of points that need to be considered.

- Something as simple as the system's name, even when carefully chosen in collaboration with the course team, can provoke fairly strong negative reactions.
- The format of messages should conform to the expectations of the users, and the 'culture' of the conferences. Confusing formats reduce the impact of the information.
- Stories provide 'a view from another angle' that can be useful to the students.

Finally, the direct (private, rather then public), messages the VP received during the study has suggested that we rethink our interaction model. Although we felt that having students ask questions in the conference helped others observe the discussion and learn from it, we feel that this has put off some students from asking questions and gaining from the information available. This is not backed up by the questionnaire responses, question six in figures 5.5 and 5.6, where a majority has said they preferred to use the conference approach. But if they did then it seems likely that more students would have asked questions, and we would not have seen those who attempted to ask questions privately. This is definitely an issue that needs to be investigated in the second study.

## 5.5 Summary: What have we learned?

The original design of the VP did not fully take into consideration the 'stakeholders'. Actually, it wasn't even clear who the stakeholders were at first. Although we had some ideas about what would work and how the students and tutors would respond, the original system was designed around a point of view perhaps skewed towards the autonomous agent approach, mixed with a little bit of 'we know best'. There were other problems, too. The students had to interact publicly with the system if they were interested in learning more, and this seems to have been off-putting for some. The tutors felt that the system was a black box over which they had no control, this made it an outside source of potential trouble.

Collaborative filtering systems were discussed in chapter 2. One feature of these systems is that they usually allow direct feedback. The VP has so far only allowed indirect feedback on its messages, i.e. people can make comments on its messages in the same conference. However the system's ability to understand this kind of message is very restricted. Incorporating a separate direct feedback mechanism would allow users to provide comments directly on any messages. This feedback mechanism could also include a simple rating scheme like a recommender system (Resnick & Varian, 1997). These ratings, along with an analysis of whether users asked questions about each possible story, could then be fed back into the knowledge base, perhaps allowing us to trim unnecessary or no longer relevant material. This same feedback may also help us to restructure the knowledge base by highlighting important or irrelevant material. It will also provide a measure of how accurate the students feel the system has been in contextualising its information.

The knowledge base could also be semi-automatically updated by identifying threads that did not trigger a match and seeing if these could be used as new stories. Those threads that did trigger a match could be used to enhance the information that is already there. The threads which incorrectly triggered a match could help us to reassess the criteria used in matching.

It seems that all the major stumbling blocks to the acceptance and use of this system are not technical but social. The social acceptance of the system needs to be addressed, by providing clearer benefits to both the tutors and the students. For the tutors to accept and make use of this system they need to be able to feel that they control it, in this way they will be in a better position to exploit the potential benefits of the system. The students currently only see the information that the VP reveals during its interactions; in keeping with the idea of FAQs we intend to make all the VP's information openly available to those who wish to browse it. We will still maintain the VP's active presence within the conference, while also making the knowledge base available in an archive conference.

## 5.6 Developments for the second prototype

The second prototype of the VP was designed to enhance the first version of the system and extend its abilities to new areas (e.g. automatic welcoming of new students, and direct tutor control). From the initial work and feedback received a number of design decisions about the operation of the system have been made. These choices have been influenced by the feedback received from the students and tutors, and the observations made during the first study.

To best understand how the VP compares with other systems and the intended effects of the changes being made to it we will return to table 2.1 and consider the VP in the terms of: interaction with users, format of cases, retrieval and maintenance. In addition to these four aspects we will also consider the VP's role as a 'Tutors Assistant'.

*Interaction with users*

The VP is still intended to interact with its users in an active way similar to that of ContactFinder. Rather than waiting for the users to approach it with questions the VP actively seeks out identifiable problems in the current discussions, identifying current

topics, and joining in opportunistically with messages. The students can then ask questions about these messages to retrieve more information.

With Uncle Bulgaria the structure of the case-base was hidden. The questionnaire feedback indicated that a number of the students would like to have access to the whole case-base. The tutors also had the same request, on the grounds that they could then, at the least, verify the contents, and perhaps add to them.

For the next presentation we will make the whole case-base available in a read only conference. This will take a form similar to that of an FAQ, which is how we continued to refer to it, in a read-only conference in the 'Archive' area on FC. This 'FAQ' form was of the 'root' message of a thread posted with its headline as the subject (i.e. the question). All the 'sub messages' were then entered as replies (i.e. the answers). The content form of the threads supported this arrangement well. There are three reasons for adopting this approach:

- the course team would like an FAQ available for the students
- it would permit passive access to all information
- it would spread the maintenance load (see the maintenance subsection below).

We also hope that the change in presentation to that of FAQs from previous tutors and students will address the earlier student comment that *"...it is clear that it is not a student or a tutor"* [appendix 5, 6] but rather a tool for presenting the experiences of their predecessors. In addition to this change of style the name will also be changed. The electronic tutors and I brainstormed some ideas for a new name during which we developed the following criteria for the name:

- It must be clear it is not a tutor or a student
- It must be clear it is not human
- It must be as neutral as possible
- The name must reflect its role

With these in mind, and considering that the case base was also to be available through the archive area, after coming up with ideas such as 'Robot Tutor' & 'Virtual Alumnus' we eventually settled on the name 'Active Archive' as fulfilling all these criteria.

*Format of cases*

In the first study, the typographical presentation of the messages did not exploit the abilities of FirstClass, or fit into the expected cultural norms of conferencing on this course. This shortcoming was due to the technical limitations of the FirstClass command line interface. SoftArc, the makers of FirstClass, provide a 'gateway' interface to enable FC servers to talk to each other. It is possible to exploit this software so that the VP can act like another server. This would enable messages to be sent containing richer formatting and styles. The second prototype used this approach to address the formatting problems the first prototype had been criticised for. This additionally permitted consistency between the style of the messages sent by the second prototype, the Active Archive, and those which were directly available to all students through the new archive area.

*Retrieval*

In the first study the only way cases could be retrieved was through the automatic identification of relevant discussions by the VP. With the second study the case-base was to be open to manual searching by the students. The automatic retrieval of the VP would have certain controls imposed upon it. The tutors would be able to monitor the current state, and make any changes they saw fit. We also discussed introducing a refractory period - a time of waiting between the triggering of a message and it actually being sent, as part of the control mechanism. The intention here is to allow the discussion to develop further before the VP joins in - this would also permit the use of an 'approval' mechanism where the tutors could be notified of a pending message and have the option to intervene. The idea of an almost instantaneous response to the student is something which is against the norm of asynchronous conferencing. With the login patterns of most tutors and students a response to any message may not be forthcoming for up to a week. And anyway, a student using the offline reader would not receive their response until they next login to check for new messages.

Secondly, the question answering system was to be changed to allow students to send direct e-mail to the VP, and although asking questions through the public conference was still supported, the students were to be discouraged from doing this. A reply to a direct e–mail would go directly to the student, but if sufficient students were interested in the same questions, follow-up messages would also be posted by the VP to the conference.

In addition to these changes we also aimed to develop better retrieval algorithms, perhaps by using a case-based natural language system similar to that described by (Cardie, 1994). This approach would treat the sentence level structure of messages as individual cases. Each discussion thread would be made up of sets of messages. The story base would in turn be made up of sets of cases for each story. This would enable more accurate matching to take place, based more on the content rather than on certain indices, as is currently the case.

*Maintenance*

Maintenance is currently in the hands of the VP's administrator. With the case-base being open to the public it would be better to allow students and tutors to submit additions to be considered. This is the most important area where there would need to be a clear benefit to the students and tutors to ensure that maintenance is actually achieved. To make sure that the submissions are academically suitable it would be necessary for a tutor to edit and approve them. There are two reasons for this, first it would help to keep the information up to date, and second, the tutors would then benefit in following years with these answers already being available. For the students, and the tutors, we rely on a certain amount of altruism for submissions which could be 'rewarded' by being personally accredited for the new cases (currently they are all deliberately anonymous). Additionally the system could help to recognise and invite possible additions to its case-base. The users would then be asked if they would mind their message being added to the FAQ list and whether they would like to contribute anything else. This would give them a chance to make a permanent contribution to the course.

*Tutor's Assistant*

The VP has always been intended to be a tutor's assistant. As we mentioned earlier there are a number of tasks that tutors do that could be augmented, or even automated, by a system like this. Our aim for doing this is to increase benefit to and motivation of the

students, and to reduce the burden on the tutors. Experiences with asynchronous conferencing have highlighted two other areas where the VP can help:

- Welcoming students the moment they get on-line.
- Structuring discussions by introducing topics at a pre-set time.

It has been found that greeting students and offering words of encouragement, even if automatically, helps improve their motivation and the likelihood that they will continue to use conferencing (Salmon & Giles, 1997). There needs to be an obvious benefit to conferencing over just having the ability to contact a wider range of people and leave messages. The content must be of a high quality and appropriate to the students as this medium has to compete with telephone and face-to-face contact. It was intended to implement much of this functionality in the second prototype.

One other role which has been suggested by students for the VP is to provide summaries of the current discussions. This would then save them the time of actually reading the discussions. There are two problems with this; first it removes any likelihood of a student using this approach actually contributing to the discussion, and second, I believe that this need is driven by the perceived (from the perspective of the students who ask for this facility) low quality of the discussions in the conferences in the first place. Providing summaries would negate the teaching use - or even any use - of electronic conferencing. The students would have no motivation to take part in the discussions, as they could just read the summaries. Of course, without discussions there can be no summaries.

The second prototype will take into account the importance of showing a clear benefit to all the users of the system. With the changes we have made we hope to observe an increase in the use of the VP by the students, and in the use by the tutors of its assistant capabilities. We hope that our next trial will be proven successful enough for the VP to be fully adopted by this and other courses.

It has been suggested that the second prototype should adopt some ideas and experiences from collaborative filtering systems, and specifically that it should be possible to measure the relevance and importance of messages based on feedback received from the user. To this end, we intend to explore the viability of this approach by incorporating a feedback mechanism into the VP.

## 5.7 Summary

This chapter has looked at the opinions of the many different kinds of users of Uncle Bulgaria. We have presented the feedback received from the questionnaire and the other sources of comments. This feedback and the other evidence received during the trial of the first prototype has allowed us to develop a range of new ideas as to the approaches that should be taken to encourage the use of this system. These points and ideas have led to the development of the second prototype, the rationale, goals, and design of which is described more fully in the next chapter.

# Chapter 6: The Active Archive: a new Virtual Participant

## 6.0 Looking forward to the Active Archive

This chapter takes the analysis of the first prototype in chapter 5 forward, giving some pointers to what it would be nice to achieve, and grounding it in the practicality of what is actually possible. It is perhaps unfortunate that some of the ideas, such as greeting each new student, fell by the wayside, but they should still be kept in mind when looking to expand and extend the current capabilities. This chapter discusses the changes that have taken place to create the Active Archive, then chapter 7 will review what actually happened in the second study, and what the students thought of it.

## 6.1 A change in focus

The original concept for the Virtual Participant was as an autonomous program which operated without any outside intervention. The aim was to minimise any human control over the system so that it could function separately without any supervision. Because of this it was designed black-box style so that the inner operation and state was hidden. This autonomy-centred approach was isolating to the tutors. It also meant that the knowledge base was effectively hidden from both the students and the tutors. A number of students (statement 9 in chapter 5, figures 5.3 and 5.4) said that they would have liked access to the complete knowledge base. The tutors have also said that they would have liked access, so they could validate the educational content and edit anything that could be misleading. These requests, and the desire to get the system more aligned with the running of the course, has suggested that we adopt a more user-centred and socially aware approach to the design of the second prototype.

When approaching the design of the second prototype, my intuition was that even if the system was opened up to allow outside control, the tutors would choose not to exercise that control. My feeling was it was likely that they would be content in the knowledge that control was there to take, even if they did not exercise their right to take it (Langer, 1983). For the students we made the other changes discussed in the previous chapter, altering the name, allowing private interactions, introducing a feedback mechanism, and providing them with access to the entire knowledge base. We also wanted to look at ways of linking the updating of the knowledge base into the work cycle of the students.

## 6.2 Changes for the student

The most important users of the Virtual Participant were the students. How the system was presented to the students, and their manner of interaction with it, were of utmost importance in encouraging their acceptance of the approach. If the VP was to gain increased acceptance in the second study compared to the first, then there needed to be some well thought out changes. In the previous chapter we have highlighted the areas which have received the most feedback, here we describe the exact changes as they affected the students:

- **Publicly accessible knowledge base:** To address the need to make the knowledge base more accessible it was decided to place it in a read-only conference. To do this a new conference was created called the 'Archive'. In this conference the entire knowledge base of the Virtual Participant was stored. As other conferences became inactive during the year the tutors also placed these in the archive area.
- **Naming the Virtual Participant:** Although the name Uncle Bulgaria had proved popular amongst quite a few of the students, and provided a metaphor for the approach, a number had been put off. A less contentious name was needed. In discussion with the computer tutors we explored the options for new names. They wanted it to be clear that the system was not a person, and also that it was not 'official dogma'. Because of this the system could not be called the 'Virtual Tutor', nor my favourite suggestion the 'Virtual Alumnus'. Eventually we settled on the 'Active Archive' as being a name that was clearly to do with the archive (see the previous point), and yet not a human, nor something that was teaching the students.
- **Personal interaction:** Feedback from the questionnaire (see chapter 5, figure 5.6 statement 6) indicated that the students were happy with the initial approach to interaction, publicly through the conference. However the lack of interaction during the first study, and the number of students who sent UB personal messages, seemed to indicate otherwise. Because of this we decided to allow the students to interact with the system by personal e-mail. They could still interact with it publicly but were not informed of this. With version 3.5 of the FirstClass client (the OU standard version during the second study) we could also move away from the copy-paste ritual necessary to ask questions in the first study. In version 3.5 it was possible to select portions of text and then simply choose the reply option. This would then include the selected text in the reply. In this way the method of asking questions was simplified

so that all the student had to do was select the question(s) they wished to ask, choose 'reply to sender', and send the message. The reply would take the VP between 20 and 40 seconds, although extra time delays could occur because of First Class.

- **'Common Interest':** So that other students did not miss out because of the personal interaction, the idea of 'common interest' was introduced. By this we meant that any question that had been asked more than X times by personal e-mail, where X was rather arbitrarily set to five, was then also posted to the coursework conference as it was considered to be of common interest.
- **Feedback:** To enable the assessment of how useful the Virtual Participant had been a feedback mechanism was provided. In just the same way that a student could ask questions, they could give feedback. At the bottom of each message sent by the VP the user was given a chance to rate how relevant the message had been to the discussion and how relevant it had been to them personally. The student or tutor could also include any text comments they wished to make. This feedback was sent just like asking a question, and it was even possible to combine the two in a single reply.
- **Delayed response:** It was not yet clear whether the immediacy of a message from the VP would stop a thread developing. To address the possibility of it being a 'thread killer' we introduced a delaying mechanism to allow the thread to develop more naturally. In this way messages from the VP would be delayed by a certain amount of time, or until a certain number of new messages had been posted to that thread, to give the students a chance to discuss the point amongst themselves first.
- **Greeting students and structuring discussions:** These two suggestions came from other work on motivating student conferencing (Salmon & Giles, 1997). In practice, the 'water-rats' were not interested in having students greeted automatically, and they were very much against having the Active Archive introduce discussions, as they felt it would interfere with their own teaching style.
- **Closing the loop:** Finally, to help feed back into the system what they had learned, as part of their revision the students were to be encouraged to write down any important lessons they had learnt from the course. These could then be incorporated into the knowledge base, so that future students might benefit.

## 6.3 Changes for the tutor

There are not so many key changes for the tutors, as they did not interact with the system to the same extent. However, they expressed a strong desire to be more aware of its operation and current state. To this end we aimed to provide the tutors with greater control over and access to the working of the system. The intention was to bring them more fully in as stakeholders, by seeking their approval of the system. The first part of this was to give them access to the knowledge base. By being able to see the entire content of the system's knowledge they could feel more comfortable with the messages it would send.

Secondly, with the delay mechanism described above, it would be possible to provide the tutors with both notification and approval mechanisms. In this way the tutors could be kept up to date with the system's likely future behaviour. It was proposed that the tutors could interact with the system either by e-mail or through the web. The VP could also keep them informed with e-mail updates, and the approval mechanism implemented through the same method. Initially this proposal was met with much enthusiasm. However as the start of the new presentation came closer anything which was not directly related to the teaching and conferencing tended to be ignored. To this end the tutors eventually decided that they were not interested in having a web interface. For this reason it was never implemented, however a monitoring page would be straightforward to included in the system.

Finally, a continuously running system could provide other facilities to the tutors. For example it would be possible for the system to personally great each new student as they got online and encourage them to conduct a few initial tasks, like filling in their resumé. The system could also be given messages to send at pre-set times for example, to encourage the students to discuss certain things at certain points in the course. These were all straightforward extensions to the new framework which, as mentioned in the previous section, the tutors were either against or not interested in.

## 6.4 The Active Archive

The second prototype was in practice a complete re-implementation of the first. The aim was to make the system more reliable, and as extensible as possible in the light of the various changes that had been suggested. It was expected that each course that would

choose to use this system would be interested in tailoring its abilities to their specific problems, so modularity and extensibility were key requirements.

The most obvious immediate change to the VP is that it used a different method of interfacing with First Class. Previously it used to connect by the command line user interface (CLUI): this had a number of disadvantages in terms of presentation and communication (for example, it prevented use of fonts or styles), but it did allow most of the functionality available through the graphical user interface. In the second prototype we moved over to using a 'gateway' provided by SoftArc. After a few modifications this gateway enabled the VP to send and receive messages to individuals and conferences within FirstClass while allowing font and style information to be retained. The gateway also allowed the VP to set message flags such as priority, sensitivity and receipt preferences so that it could keep track of its messages. A most useful addition is that messages passed through the gateway contained additional information not available to the normal user, that allowed the VP to track threads of discussion more easily. The role of the CLUI was not eliminated, however. If the VP needed to access history data, directory data, read resumés, confirm who is on line, or use the 'chat' interface, these still had to be done through the CLUI.

It was not actually necessary to use the history information in the second prototype. There were several reasons for this. First the gateway interface provided threading information much more directly. Secondly, with the Active Archive we opted for matching on complete threads rather than discussion branches within threads. This also reduced the need for history information. However, to allow for detailed analysis of who read which messages, and to select the candidates to be surveyed, certain history information was saved manually. This was separate to the running of the VP.

The second fundamental change to the VP was that it ran continuously in the second study. The system would poll the gateway approximately every 20 seconds and then deal with the messages it received. Initially the system ran constantly. However to reduce load on the server during the night (when there was little student usage and the backups and updates occurred) the VP was later set to run only between 6am and 1am every day.

These steps were all directed towards making the VP more flexible and extensible. It was never intended to be 'one style fits all'; different courses would wish to use it in different ways. To this end we tried to incorporate as many options as possible so that with more

enthusiastic users the VP could take on a more comprehensive role. However, in this study, I will focus on the key changes I have just described, and on the effects they had on the acceptance of the Virtual Participant.

## 6.5 A new system design

Back in chapter 3, section 3.4, we looked at the system description for the first prototype. In this section we will take the same approach for the second prototype. Because of this these sections bear certain similarities which are useful in highlighting the changes. Not all the proposed features were implemented, and this is indicated where necessary. The reason behind most features being left out was a combination of a lack of time and a lack of interest in that feature. Specifically neither the web based interface nor on-line monitoring and greeting were implemented. The chat interface has also not been added to the VP so far, and could make an interesting future project.

Like chapter 3, this section uses various components from the KADS system analysis and design methodology (Tansley & Hayball, 1993) to give a clearer picture of the function and operation of the second prototype, and to enable direct comparison with the design of the first. Again we begin by developing a system-overview document. In the case of the second prototype our aims have clearly moved on from a proof of concept to demonstrating a fully functional and successful system.

*Objectives of Prospective System*

As before, let us consider the various "agents" within the system. Based on last year's study we were able to make an assessment of how important the various members are. Those who are no longer considered directly relevant are shown in italics:

- Tutors
  - *Normal Course tutors*
  - Electronic Tutors
- Students (could be broken down along many axes)
- *The course team*
- The Virtual Participant
- Myself

The normal course tutors had not made any input into the system, and had never given any feedback. The course team appeared to be content to allow the research to continue,

and also, at this time, had their hands full with the course rewrite (B882 was replaced by B822, "Creativity, Innovation and Change", soon after the end of this study). This year my goals were slightly different. As there had already been a successful proof of concept this year I was looking to show that the changes made were successful at improving the system, and that a wider number of students would use it.

So for this second year I was looking to achieve:

- Greater interaction between the Virtual Participant and the students. More interaction would increase the exposure of the students to the VP, making it easier to assess its effectiveness.
- Feedback from the students showing greater acceptance of the idea and the system. Fundamentally this was the only way to demonstrate its effectiveness.

If neither of the above goals could be achieved there at least needed to be lessons learned, which could then help to inform subsequent and other systems of this nature. As always I was looking to identify areas in which the VP could be improved.

So to summarise, the high-level aims were to show an improvement on the system in first study, and point the way to a tool which could be used on a wider number of courses. If these aims could not be achieved, the reasons for failure needed to be identified and possible solutions found.

*System Functions*

It is intended that the second prototype will perform the following functions (new functions are indicated in bold, italics indicate disabled or not fully implemented for original and new functions respectively).

FirstClass gateway system:

- **Upload messages**
- **Download messages**

FirstClass Command Line User Interface (CLUI)

- *Access to history data*[1]
- ***Access to 'chat' interface***[2]
- ***Access to 'directory' information***[3]

Virtual Participant:

- Recording messages *and history of messages* from FirstClass
- Storing a database of past cases
- Storing the current state of the system
- Matching current FC discussions to past cases
- Posting of 'matches' with past cases to FC
- ***Delaying the posting of messages for an appropriate period***[4]
- Identifying, and dealing with, public requests for more information
- **Identifying, and dealing with, private requests for more information**
- ***Greeting new users***[5]
- ***Web-based interface to current system state and administration***[6]
- ***Tutor notification and control***[7]

*Provisional System Structure*

The high-level schematic has not really changed much from that for UB, shown in figure 3.12. The majority of the changes were in the user interface where we were no longer exclusively using the CLUI, but also the FC Gateway system and the Web. The details of

---

[1]Not reimplemented in this version
[2]Never implemented
[3]Never implemented
[4]Implemented, but disabled due to change in conference management
[5]Never implemented
[6]Partially implemented, not available in second prototype
[7]Notification implemented and disabled, control never implemented

the problem solver changed, but the system processes remain the same. The same is true for the internal storage.

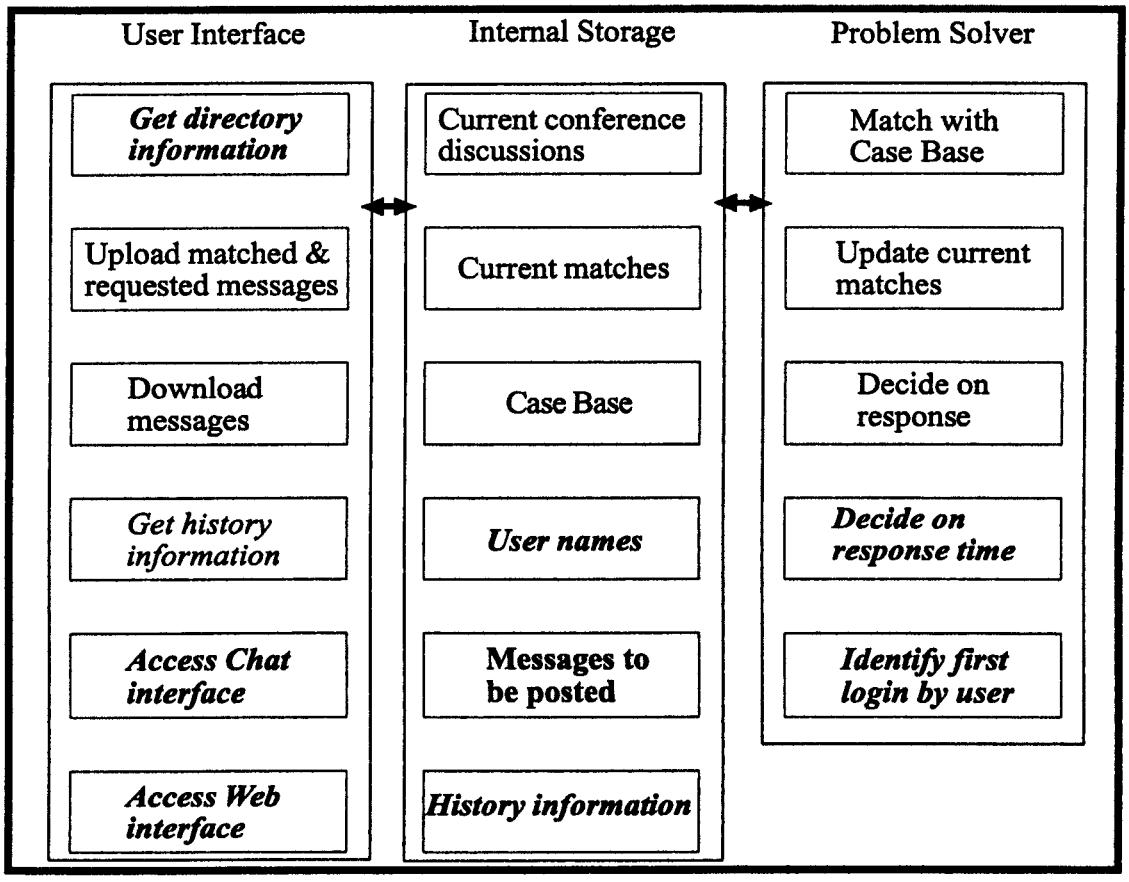


Figure 6.1: Schematic diagram of provisional system structure.

Figure 6.1 presents the provisional system structure. Not all features were implemented and certain stored information proved irrelevant. These features were not directly relevant to the functioning of the system and mainly provided useful extras. For example, the system had been intended to provide the functions needed to investigate the effect of automatically welcoming every new user on the use of the system, especially as it has been suggested that this approach could improve take-up of conferencing (Salmon & Giles, 1996).

In the user interface the ability to get history information, access the chat interface, and access the web were all either disabled or not fully implemented. History information proved irrelevant to the functioning of the second prototype system as data tokens attached to each message by FirstClass indicated which threads they belonged to. These

tokens are not visible to normal users, but are accessible as extra information passed from FirstClass via the gateway interface. However, history information can be useful when analysing the flow of discussion in more depth, especially if the offline reader has *not* been used. Because these features were not present in the interface it was not necessary to store history information. Also, the tutors changed the way that conferences were administered this year so that messages were more rapidly moved into subfolders which the gateway could not access. Because I had no control over this, and there was a risk of not being able to deal with this technically, the time delay before sending a message was set to always be zero.

*Provisional Information Requirements*

There is a lot of overlap between the information requirements specified for the second prototype compared to the first. Specifically, no changes were made to the format of the case-base and the previous information requirements are included here for completeness. Other information requirements have been redistributed, or are no longer necessary. For each of the following additions or changes are in bold, disabled functions in italics.

In the case-base we store (as in Figure 3.8):

- 'Anonymised' messages from previous discussions.
- These are arranged as threads.
- Threads are clustered into topics.
- Topics have been grouped into stories by manual editing.
- A data structure representing each of the stories, the matching threshold, and a current match level of 0.

In the thread database we store:

- Messages posted to FC.
- These are arranged as threads **using their identifying tokens.**
- *The **full** history information for each message.*

- On a thread by thread basis we store the current state of matches with the case base. **This is maintained for the whole thread rather than on a message by message basis.** When a message is received which pushes the match value over the threshold a reply is triggered to that message.

In the user ID database, new in the second prototype, we store:

- ***The user names and ID's.***
- ***If they have logged in yet, and if so how often.***

This concludes the system overview document. We will now look at the data flow diagrams for the second prototype.

*Process analysis: Data flows*

The following data flow diagrams show the process that the VP goes through in dealing with new messages. This is slightly different to the process used in the first prototype and these diagrams should help to highlight the differences. Input from history information is shown where it would have been used, although it proved unnecessary for the operation of the system. The delay in sending messages is also shown although it was disabled in the second prototype.

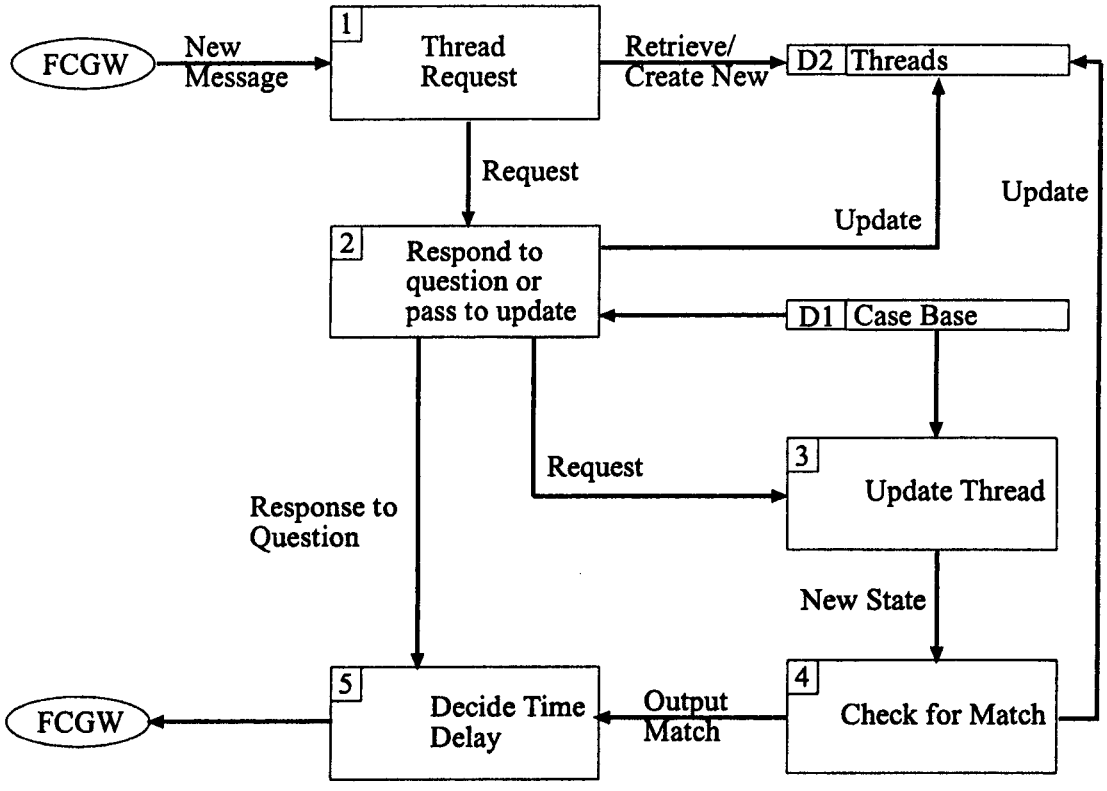


Figure 6.2: High-level data flow diagram for second prototype.

The data flow diagram presented in figure 6.2 provides a schematic of the data flows and processes involved in the second prototype. These processes represent the following:

1. A new message comes in to the system from the FirstClass Gateway (FCGW). This causes the system to generate a request to retrieve the relevant thread from the database. If no relevant thread is found a new thread is returned.
2. The retrieved thread and the new message are then passed to the question identification stage. Any question identified in the message is retrieved from the case base and sent to process 5. If no question is identified the message and thread are passed to the updating process.
3. The updating process uses the new message to update the values for matches between the thread and the case base. This generates a new thread state which is passed to the matching process.

4. The matching process identifies any case in the case base which matches with the thread above its threshold value. If a match is identified then it is passed to process 5. Either way the thread database is updated with the new state.
5. This process implements the waiting period between a message being queued and when it is sent to the FCGW. In the study this period was sent to zero and all messages were sent immediately.

Processes 1 and 2 break down to produce slightly more complex data flows as shown below:

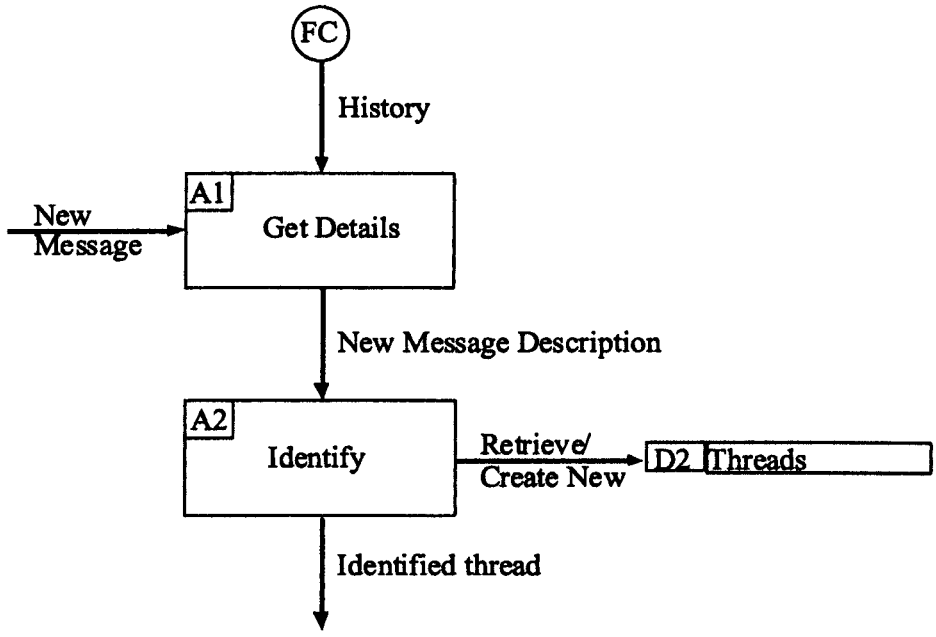


Figure 6.3: The Data Flow Diagram for the Thread Request process.

In figure 6.3 the input from the history data of FirstClass is shown. This data proved unnecessary in the normal operation of the second prototype and is only shown here for completeness.

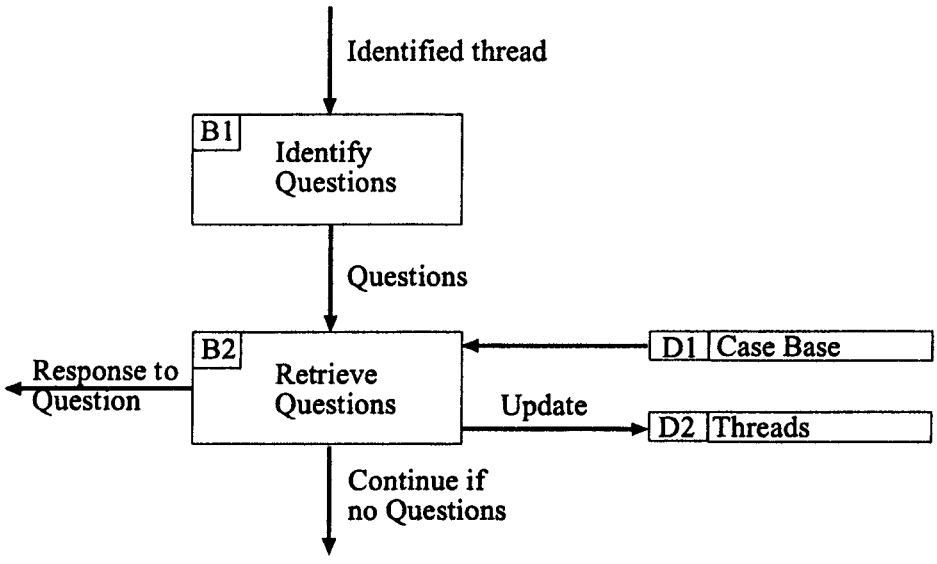


Figure 6.4: The Data Flow Diagram for the Respond to Question or Pass to the Update Thread process.

In figure 6.4 the details of the question identification stage can be seen. The first stage is to identify any questions present. Those questions are then retrieved and sent as appropriate. Otherwise the message is passed on to the 'Update Thread' process.

This section has shown the data flows involved in the second prototype of the Virtual Participant. The next section returns to the selected Generic Task Model, and again compares the first and second prototypes.

*The Generic Task Model.*

With the second prototype there was no change in the generic task model selected for this process. Once again, the VP uses the "System Analysis - Identification - Correlation - Assessment" GTM from the library in (Tansley & Hayball, 1993). Also the input/output description has not changed from that used previously. This description is shown in figure 3.13 and repeated here in figure 6.5.

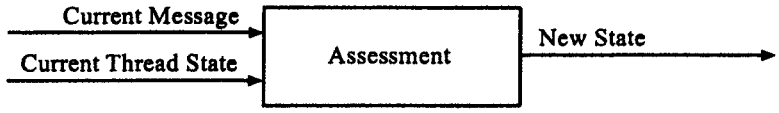


Figure 6.5: Input/output descriptions for the VP.

Other possible inputs to the system include that available from the directory information, and from the web-based interface. These are not relevant for the expertise part of the system.

In the second prototype each message coming in to the assessment system is a message coming from the FirstClass gateway for processing. The matching process now treats each thread as a single entity rather than watching for various topics to emerge on different branches. Therefore, the current thread state is the value for matches between the case-base and the current thread in total, rather than just at this point in the thread as in the first prototype. If the message starts a new thread the current value of all possible matches with stories is set to zero. The new state produced is associated with the whole thread, not just the current message in the thread.

As before, we have adapted the Generic Task Model (GTM) for System Analysis - Identification - Correlation - Assessment to reflect the inference processes going on within the Active Archive when dealing with messages. This is shown in figure 6.6.

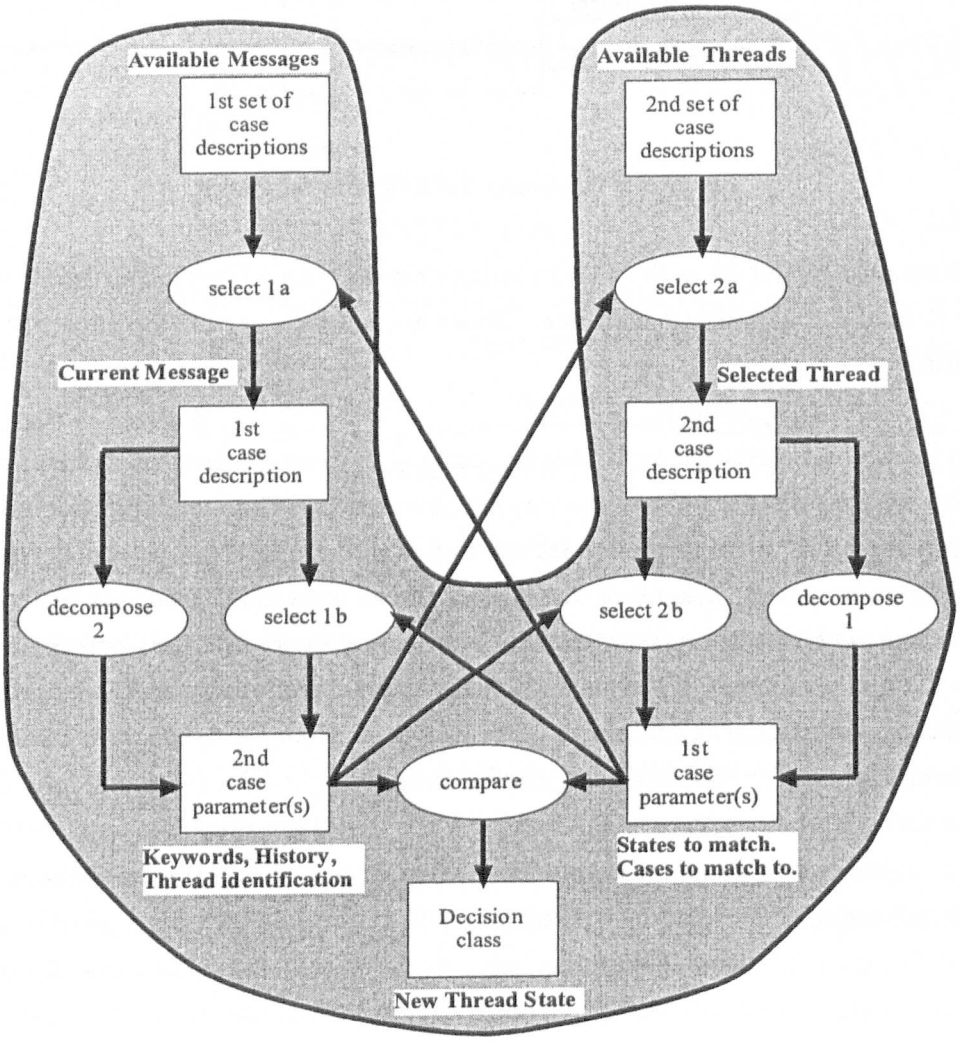


Figure 6.6. The Task Model for The Active Archive.

Unlike Uncle Bulgaria the Active Archive uses the whole of the GTM, including 'select 1b' and 'select 2b', which were previously not used. The new task model, shown in figure 6.6, shows the various stages of the system's operation. The incoming messages come from the gateway and each is dealt with in turn on the basis of the oldest first. From each message the unique thread code is used in the selection (or generation) of the thread it belongs to. The message is also decomposed into its keyword information. The current thread provides information about the previous messages, about matches which have been sent to the conference, and the current state of all other matches. The compare task uses all the available information to either post a reply, or to update the current thread match state and post any match above the threshold.

This section has given an overview of the processes performed by the second prototype and how they differed from those in the first prototype. In this instance the goals of the system were more clearly focused on improving the concept and demonstrating the success of the revisions. The next section reviews what happened when introducing the second prototype in comparison to the first one. The next chapter then reviews the interactions between the users and the system, and discusses the feedback received from them.

## 6.6 History of introducing the Active Archive

The second year trial with B882 was the first time when the majority of students would already have been familiar with FirstClass, especially those having just progressed from the compulsory MBA Strategy course, B820. Also, those students who had already been registered for conferencing during a previous course were allowed to keep their accounts, so a number were on-line early. Because of this it was decided to start the conferencing one week early on the 9th of February 1998, even though it had originally been planned to start on the 16th. A message about the Active Archive was placed in the coursework and welcome conferences at this time. At this point the archived TMA1 messages which the AA held in its knowledge base were available in the archive conference. Due to a number of delays, and after extensive testing, the Active Archive itself was finally started on the 11th of March.

This year the conferencing structure was slightly different. The tutors abandoned the structure of separate conferences labelled TMA1 and Block1 under the coursework conference. Instead they only had the TMA1 conference, and the coursework conference was no longer read-only. This meant that the students could post to either, and so could the AA.

Also this year, the thresholds were lowered somewhat, as in the first year they had been set deliberately high to keep the tutors happy. This had meant that previously some messages had not been triggered in situations where they were relevant. Although for this year we had built in certain mechanisms for delaying the posting of AA messages, because the tutors had decided that after a while threads of discussion would get archived to 'folders' (which could not be accessed by the AA or the off-line reader) this was abandoned.

Coming in late (although there were still a number of students who had not started conferencing either) the AA already had a number of messages, 7 in total, to post on the first day. Within 2 hours students were asking it questions, 6 different ones by the end of the day. One of these questions was asked publicly in the conference. That evening the student's message asking the question was deleted, but the AA responses were left in the conference.[8]

Within a few days we reached the point where one question had been asked 5 times and so its answer was to be posted to the coursework conference.[9] Initially 'common interest' messages like this were cross-posted between the TMA and coursework conference as the AA usually couldn't identify where the original discussion thread was. Later on when more 'common interest' stories were posted it was decided that they should only go to the TMA conference. By the end of TMA1 the AA had posted 8 unsolicited messages and 7 'common interest' messages.

Soon after this we started with TMA2. The TMA1 conference was made read-only and moved to the archive. The TMA2 archived messages available to the AA had already been available to the students a few weeks after the AA started in TMA1. At the beginning of TMA2 the AA was rather enthusiastic, because of the lowering of the thresholds. We faced the problem of too many AA messages in one particular thread. Initially I was inclined to let it continue without any intervention. One of the tutors felt that because some of the messages were not relevant to the thread then we should change their subjects, but in the end we agreed to let it ride to see if we received any negative student comments. After it had posted a number of messages I manually altered the threshold values stored for that particular thread to help suppress further messages. After I had done this I went through the thread to reconstruct the details of what had happened.

---

[8]During this day the AA unexpectedly crashed because students were sending it messages with no body. At the time it seemed most likely that this was just a technical glitch from students using version 2.6 of the client. Later, it became apparent that the problem was that students weren't clear about how to ask questions and in practice the AA should have been responding to these messages, and providing help.

[9]This caused the only major bug where the AA repeatedly posted this message. The problem was quickly resolved and the excess messages deleted.

It turned out that the first two messages in this thread were quite long and of great depth. Because of this they touched on a number of different topics in Block 2, triggering several matches. These matches were stored in an agenda, but the AA only took the top one in response to any message. This meant that every time someone posted a follow-up message the AA posted another matched topic. I had not foreseen this situation and it clearly shows the need for more advanced heuristics in the management of message posting, as well as a better system for matching.

After suppressing these responses, but without making any other changes (for example, to the message subject) the AA was allowed to proceed as before. Even though the messages were not always directly relevant to the current discussion they did generate a number of questions from the students. By the end of TMA 2 the AA had posted 15 more unsolicited messages and 13 'common interest' messages. During TMA2 the tutors set up another coursework conference called 'Coursework Specials'. I had no access to this conference, which included special conferences on certain specific topics and several collaborative problem solving conferences. It is likely that the AA had some information that could have helped in these conferences, but it may be that it would have needed to be more specialised.

Towards the end of TMA2 I began to design the questionnaire that was to be sent to students to help evaluate the second prototype. A few days after the submission deadline for TMA2 the AA was shut down and the questionnaire posted out. A more detailed analysis of what went on in this second study, and the students' responses to the questionnaire, will be discussed in the next chapter.

## 6.7 Summary

In this chapter we have described the design changes which were planned for the second prototype VP in the light of the first prototype. We then described the history of the second study. Not all the planned changes were made, and some were removed or disabled. However it is clear that a much tighter integration between the use of the tool and the teaching must be achieved, and it can not simply be an extra tool. The next chapter looks in more detail at the interactions which took place with the Active Archive, and the students' reactions to it.

# Chapter 7: Interacting with The Active Archive

## 7.0 Overview

> "In this section we wish to present some trends and observations of the uses, strengths, and weaknesses of this ... tool. ...In this section we have tried to be as candid as possible about the weaknesses and research problems... We hope this candour does not create an overly negative impression about what we feel is a very positive research effort." (Conklin & Begeman, 1988, p. 147).

This chapter will look at how students interacted with and perceived the second prototype. The changes to the system described in the previous chapter were intended to address the shortcomings of Uncle Bulgaria. The second prototype and study provided the opportunity to test out these changes. Continuing the approach adopted for the first prototype we will look first at the use made of the system in this second year. We will then move on to look at the patterns of interaction and the feedback received from the students.

## 7.1 How much use was made of the Active Archive?

First of all, we will compare the number of registered students who interacted with the Active Archive with the number who interacted with Uncle Bulgaria in the previous study, shown in section 4.4:

| | Active Archive | Uncle Bulgaria |
|---|---|---|
| Number of registered students | 834 | 618 |
| Number logging in at least once | 653 | 556 |
| Read 1 or more VP message | 402 | 363 |

Table 7.1: Numbers of students involved in both studies.

Conducting a set of 2 by 2 chi-square tests on the data in table 7.1: Registered vs Logging in (not significant, $p < 0.1$); Registered vs Reading 1 (not significant, $p < 0.5$); Logging in vs Reading (not significant, $p > 0.1$). Those logging in indicate the total potential audience, and those reading at least one message indicate the number of students exposed to the VP. The results indicate most importantly that there was no significant difference between the studies, in terms of those logging in and the chance that the students read at least one VP message. This indicates that the VP did not significantly increase the amount of participation, because if there had been more participation a greater proportion of students would have read at least one VP message.

The Active Archive posted many more messages to the discussion conferences than Uncle Bulgaria, with many questions being asked and leading to the posting of common interest messages. The AA posted 42 messages to 11 threads, UB posted 17 messages to 7 threads. A 1 by 2 chi-square test indicated that the change in the number of messages was significant ($p < 0.01$), but the change in number of threads was not significant.

Figure 7.1 shows in detail where the messages were received and sent. The personal category consisted of messages from students asking questions by direct e-mail, providing feedback, or both. More replies were sent than messages asking questions received because students were able to ask more than one question per message. The administration category consisted of test messages from myself to confirm the functioning of the AA, and messages which the AA received which it could not deal with for any reason and were therefore forwarded to me. The remaining three categories refer to the discussion conferences. B882 coursework was active for the whole length of the study. Both of the other conferences were active only for their respective TMAs. Of the messages posted to the TMA1 and Coursework conferences three were cross-posted by the Active Archive as they were questions which had been asked at least five times (these three messages are shown in both the TMA1 and the Coursework columns in figure 7.1). This practice was discontinued towards the end of TMA1, at the request of the tutors, with this type of message only being posted to the TMA conferences.

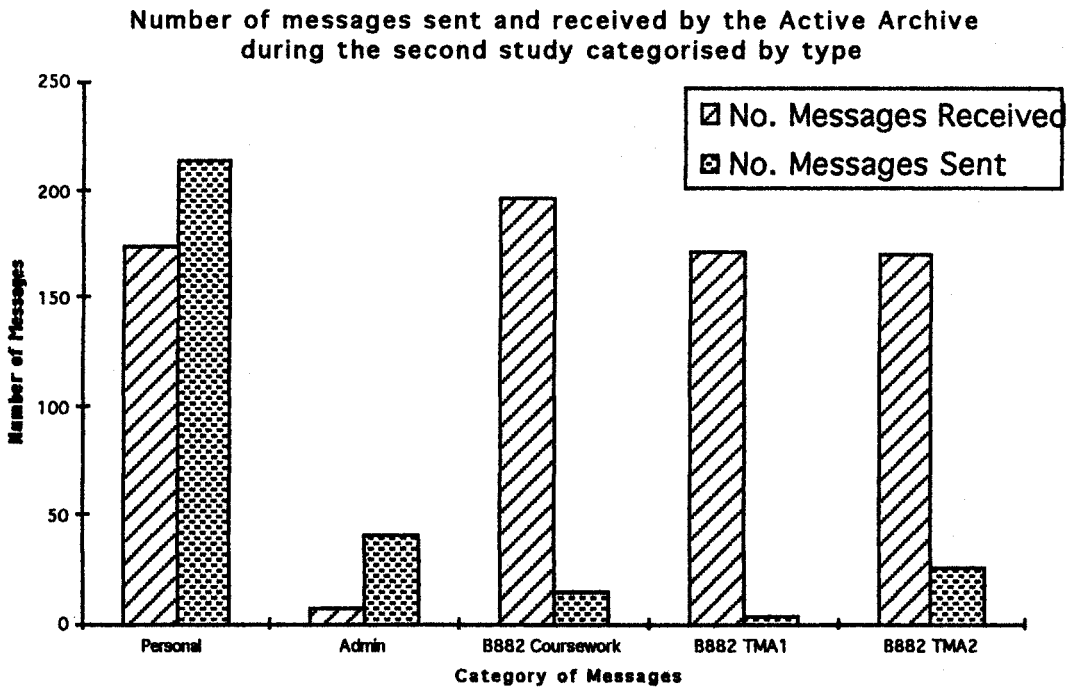


Figure 7.1: Where the messages received by the Active Archive came from, and where the replies were sent to.

Having already identified that there was no significant increase in participation this year, although there was a significant increase in the amount of interaction from the VP system, I will now look to see if there was any effect on lurkers. Appendix 6 contains a breakdown of the different types of users found in the Coursework, TMA1, and TMA2 conferences. The main observations to bring from this analysis are that only 105 students actually participated (by posting at least one message in the discussions) and that 40 of those 105 asked questions of the AA. Unfortunately no other students asked the Active Archive any questions, so no lurker ever asked a question, and no one who was not already contributing asked a question. However there was one example of a question being asked in the public conference, probably by accident, where the student never contributed to any of the discussions. If this had not been a public question then it would be an example of a lurker asking a question of the AA.

Later in this chapter there are two specific examples where messages from the AA actually stimulated more discussion through responses to its messages, figures 7.6 and 7.7. In the first of these the student who responded to the message from the AA had never contributed to the discussion before, although they had previously asked questions of the AA. This provides one instance where someone was encouraged to participate by the presence of the Virtual Participant - as well as two instances of extra discussion implying additional learning. In both cases the tutors chose to provide additional information, either as part of the discussion or in a separate thread.

Because I was only allowed access to the discussion conferences I can not definitely say whether the student who was encouraged to participate in the public discussions had already contributed to other private discussions. In addition, because of insufficient data, I can not clearly say whether there were some students who read AA messages or asked the AA questions before they started to participate - and even if I could it would be impossible to say whether there was any causal relationship between participation and AA messages. There were also no examples where the contents of AA messages were explicitly discussed or referenced by student messages which were not direct replies, either in the same thread, or in different threads. However there was an example of extra work from the tutors, mentioned above, where they chose to provide additional information about the content of one message in a separate thread, shown later in figure 7.6.

From these observations it is clear that there are some parallels to Ackerman's study, (Ackerman, 1994), and the three user types he identified. We have the participants, Ackerman's 'heavy users', who also asked the AA questions; the 'lurkers', Ackerman's 'intermittents', who didn't participate in the discussion, but did read AA messages (our "vicarious learners"); and the remaining students are 'browsers', Ackerman's 'tyre kickers', who were just sampling and looking for something to attract their attention.

One other thing which needed to be looked at was whether there was any relation between the number of AA messages read by a student and the number of questions they asked. The details of this can be found in appendix 6, but suffice to say there was no correlation.

Without access to detailed logging data about the use of FirstClass it is impossible to analyse fully what different kinds of students we have and what their behaviours were. However the data we have already considered, the observed interactions coming next, and the questionnaire feedback later in this chapter give a more detailed view of what the users thought than in the first study. Already we know that only those students who contributed to the conferences were willing to ask the AA questions; this unfortunately shows that the AA was only *directly* helping those who were already benefiting from the conference discussions. However there is no real measure of how much *vicarious* help the other students received from the VP's public messages.

An interesting observation of the use of the AA messages in the archive was that a number of students attempted to ask questions of the archived messages themselves - these came to me personally. In total I received 14 messages from 11 different

students. The answers they were looking for were in the Archive, and I used these in my replies to them. This indicates that not only do the students want access to the complete knowledge base, but also that the VP needs to support questions being asked directly from the archives.

## 7.2 The Structure of the discussion threads

As with section 4.5 in chapter 4, I will take this opportunity to look back over the structure of the discussion threads the VP was involved in during the second study. In the 1997 study Uncle Bulgaria contributed to 7 discussion threads in total. During 1998 the Active Archive contributed to 11 discussion threads in total. Table 7.2 shows the number of messages posted by the AA to the various threads. The numbers in brackets for the thread "Thoughts on readings 14 and 15" are because one student asked three questions in public which were then posted. Later on these messages reached the state of 'common interest', but the VP software never posts messages a second time.

| Subject | Number of unsolicited messages | Number of 'common interest' messages |
|---|---|---|
| During the first TMA | | |
| copious activity -> more on MBTI | 1 | 1 |
| Personal characteristics and creativity | 1 | 1 |
| Thoughts on readings 14 and 15 | 1, (4) | (3) |
| Reconciling MBTI & KAI Scores | 1 | 0 |
| Creativity or Relativity | 1 | 0 |
| Moved, was Intro to NLP | 1 | 1 |
| SET | 1 | 0 |
| Entrepreneurial, Champion and Creative leadership traits... | 1 | 0 |
| During the second TMA | | |
| Team building | 9 | 8 |
| Problem Solving | 4 | 3 |
| Reader, Readings, Readings 2, | 2 | 2 |

Table 7.2: Number of unsolicited and 'common interest' messages for each thread the AA contributed to.

During the time of the first TMA 61 separate threads were identified, of which the AA posted to eight. During the time of the second TMA 52 separate threads were identified, of which the AA posted to three. Comparing these values to last year's number of threads, 36 and 85 for TMA1 and TMA2 respectively, there seems to be

further evidence for a shift in conferencing patterns assisted by removing the need for students to re-register for conferencing for each course. In previous years it has been common for there to be a slow build up during the first TMA as students slowly get on-line, which is then boosted by the residential school making the second TMA quite busy. Students in this study, not having had to re-register, were already on-line early enough that, as mentioned before, the tutors started conferencing early. There was some very enthusiastic conferencing early on, which settled down in the second TMA, a very different pattern to previous years.

In the first study the students saw a total of 17 messages to 7 threads from Uncle Bulgaria out of a total of 121 threads. During the second study the number of messages posted to the conferences was 42 messages to 11 threads from the Active Archive out of a total of 113 threads. This is pretty much double the exposure in terms of messages sent, and significant ($p < 0.01$) as shown above, but messages were not sent to that many more threads, not significant ($p > 0.1$). But as you can see from table 7.2, some threads were quite strongly influenced by the AA.

*Thread Diagrams*

In the following subsections I will show some of the diagrams for the threads in the second study. Those not included here can be found in appendix 7. These diagrams follow the same rules developed and used in the first study seen in chapter 4. The letters which were used to represent individual tutors in the first study represent the same tutors this time. A slight modification occurred with increased use of the offline reader, and the use of the gateway, making it very difficult to tell which message came first if several were posted by the same person. In this case the messages are all shown at the same level. One extra item has been added, that of a box with a cross through it. This represents the case where enough students have asked a question (mainly by personal messages) that it has been deemed of 'common interest' and posted to the thread. These have only been included when it is relevant to the development of the thread, otherwise the 'common interest' messages are noted but not shown.

### 7.2.1 First TMA

First there are four threads which are worth discussing right at the start. The first is "Reconciling MBTI & KAI Scores", the text of which can be found in section 1 of appendix 7. This thread was just two messages long, the first from a student worried about their KAI and MBTI scores, the second from the Active Archive retrieving a message from a previous student who had also been worried about their KAI score.

This was a reasonable match and the student also asked the AA one of the available questions. Three other students asked personal questions and one used the feedback to say that the relevance to the discussion was excellent and the relevance to them was good. This is quite a nice example of the VP at work.

On the other hand, the second thread is quite a poor example of the VP at work. The thread "copious activity", which became "more on MBTI", is shown in figure 7.2. In this case this was a thread from near the start of the first TMA. The VP's message was last chronologically because all discussion took place before the software was finally ready. What happened was that the VP made a match with the introductory message from a tutor talking about the subjects to be covered in the conference. This was a typical failing of the VP system as it had no language understanding capability and could not resolve subtleties like this. An interesting event then occurred as a student replied to the tutor, changed the subject, and continued a completely different conversation. There were a number of questions asked, and feedback given rated its relevance to the discussion as quite low. The 'common interest' threshold was reached on one question which was then posted to the conference (not shown in the diagram).

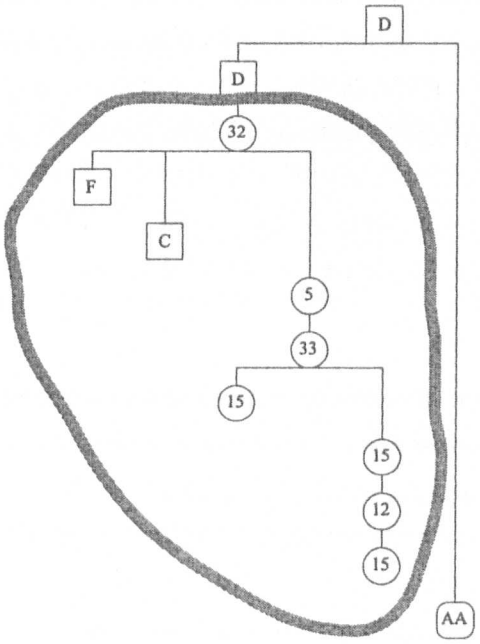


Figure 7.2: The thread "copious activity". The highlighted section is what became known as "more on MBTI".

The third thread worth describing was known as "Intro to NLP". When the AA matched this thread the message it posted was deemed by the tutors to be more relevant to the 'project' conference (the fourth TMA) and they started this conference early and

moved the message into it. At this point it became titled "Moved, was Intro to NLP". There was enough interest from the students that one of the possible questions reached the 'common interest' threshold. Some students also provided feedback, for example one asked a question then said:

"Thanks for the quick response. I assume this is a mechanised, and interactive version of the Simon Masterton postings in B882 archive, or are you giving additional advice?

As a learning tool, I think it has its merits. Too bad I'd read all the Archive material."

The fourth thread I will describe is "Thoughts on readings 14 and 15". This discussion happened early on in the course, before the Active Archive had come online. When the AA started some messages had already been archived by the tutors. The thread was a discussion about how to get started on problems, the AA matched a message on how to get started on your TMA. In this instance a student then asked follow-up questions in the conference, and the VP responded publicly. In the second study the students were not told that they could do this. It seems likely that the student's posting to the conference was accidental, because the message asking the questions was soon deleted, though whether by the student themselves or by a tutor I do not know. This was the only public posting by this student, they did not contribute to any other discussions but did they continue to ask the VP questions. These questions were of some interest, being on ways of doing your TMA, and continued to be asked by other students so that they later reached the 'common interest' threshold; however the VP never posts the same message twice. Figure 7.3 shows this thread diagram.

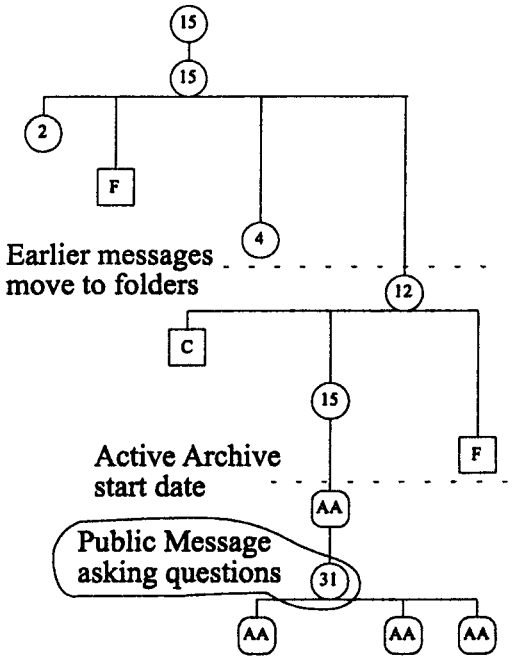


Figure 7.3: The thread "Thoughts on readings 14 and 15" showing the public response by the AA to the public asking of questions.

Of the other threads from TMA1 a few have been placed in appendix 7, and I recommend referring there to gain a fuller overall picture. The ones described here are those that seem most interesting in showing the wide range of interactions and competence of the second prototype.

### 7.2.2 Second TMA

So far we have only been looking at the threads from the first TMA. The distinguishing feature of the second TMA was that although the AA posted to fewer threads, it posted many more messages to those threads. This large amount of messages from the AA provoked some concern from the tutors. However the students made no comments and many questions were asked, and for every thread the 'common interest' threshold was reached for some of the questions. The thread "Reader, Readings, Readings 2," is included in appendix 7 as it is not particularly complex. Both "Team building" and "Problem Solving" are shown below as they are also helpful in assessing the impact of the AA in terms of stimulating more discussion and increasing the work load of the tutors.

*Team Building*

The thread "Team building", shown in figure 7.4, had 9 messages from the VP posted to it, and these had so many questions asked of them that a further 8 'common interest' messages were posted. The thread starts with a student who is interested in encouraging their "colleagues to be more co-operative and engage more readily in team work". Others mention that their organisations are similar in behaviour. The first VP message is advice on techniques for brainstorming, especially when you are not regularly physically co-located. Not a particularly auspicious start this time round. Continuing their discussion the next student talks about how this is quite common in organisations and that they should attempt to find some books on the subject. This is greeted by the AA retrieving a message on "The Ultimate Problem Solver" - a piece of humour from a previous year with no further questions. At this point the first message from the AA was beginning to provoke students to ask questions. Student 21 talks about how difficult team building is, getting people together, and getting them to generate ideas for working more effectively together. The AA follows this with its third message which is about experiences with the 'Delphi' method, which is designed to be used with groups. Student 16 then posts a message about how in one of the 'readers' the problem featured was solved by the manager by maintaining team support and shared focus. The AA follows this with its fourth message which is about creating and maintaining a shared vision in a pharmaceutical company. Student 22 replies to the problem described in this message from the AA giving more information about a similar problem they are experiencing in their pharmaceutical company - an interesting event showing that there are students in different years with very similar problems working in the same market sector. This is the first example of a message from the Virtual Participant stimulating more discussion, and a clear example of cross year relevance of problems. Student 22's message described the problems they were having with their projects and their failure, caused by the US parent of the subsidiary they worked for not considering them in the decision making process. This message was also the first contribution to the public discussion conferences by student 22, who had previously asked questions of the AA privately, showing that the AA could stimulate participation. This student later went on to start a new thread seeking input into another part of their problem, which no one followed up. The AA followed up the message from student 22, with its fifth contribution to this thread, by replying with a message about doing stakeholder analysis to define a problem more clearly. Following this branch down the two other student responses were other ideas to help student 22.

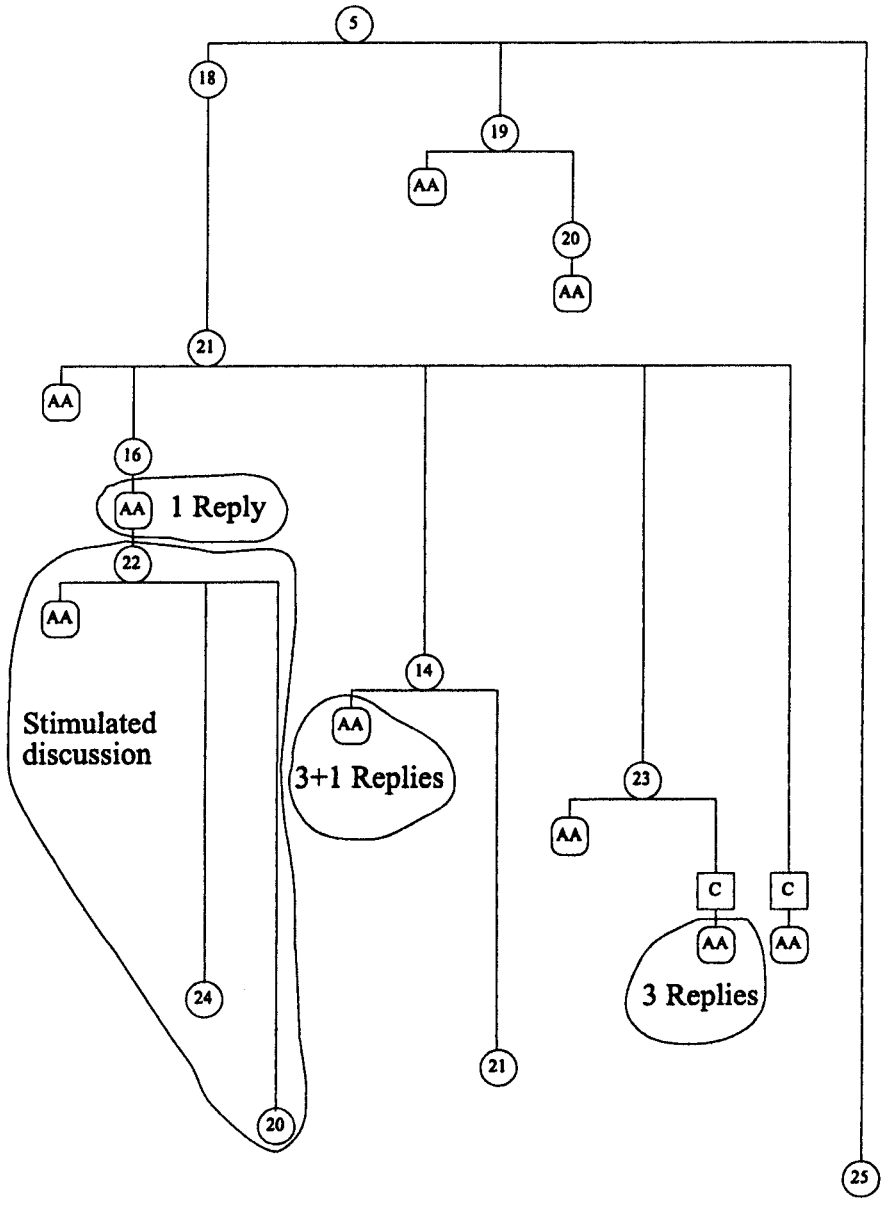


Figure 7.4: Thread diagram for "Team building".

To give you some idea of the content of these discussions two extracts from the stimulated discussion are included here, the full text can be found in appendix 8. The first extract is from the AA message student 22 replied to.

```
COMMENT: Who's Vision is it anyway

I guess I have a similar problem in that we are a subsidiary of a
company that manufactures the products we sell.  The parents focus
is from a manufacturing viewpoint and toward their local market.
What vision  can we have of our own, when we are so constrained ie
can you lead a market if you haven't control of the products you
sell.
```

The vision has to fit and be realistic. Doesn't mean you can't have longer term changes in view but the vision has to fit now if it's going to be credible.

I guess you have to incorporate what you do now, with what you could offer to your 'customers' in the future. You also have to find out what everyone else thinks assuming you want it publicly adopted rather than for your personal use.

My organisation has a mission statement and set of objectives which are used locally. It incorporates our aims to inlicence products and move further into our market place. I would be happier if we published it to our customers and to our parent company.

My organisation is a pharmaceutical company.

The second extract is the actual reply from student 22.

I have a similar problem as the small European part of a large but very US focused pharmaceutical company.

From our perspective the US want to control everything - so what is new - seem jealous of our abilities and have no interest in supporting the growth of a successful European organisation, in fact seem to us to be positively trying to stop it.

From their perspective, we are probably upstarts from these strange little European countries who mess up their perfect organisation and keep complaining that they 'do not consider the world outside the USA might have an opinion'! If it is OK in the USA it must be OK in little Europe?!!

There seems to little understanding of the economics of pharmaceuticals today, that if Europe does not contribute to sales, the NPV of most projects is very weak if not negative.

I have the task of trying to persuade them to include European staff in development teams and consider the ex-US markets and needs from the start. This is my TMA 02 planned subject but at present it seems insurmountable.

Has anyone been through this themselves and can offer any unusual suggestions about how to kickstart this process, especially with teams who are very reluctant.

Clearly these students had two very similar problems they were trying to solve. If it had not been for the VP student 22 would have no idea that such a person existed.

In parallel to the stimulated discussion student 14 follows up on the comments by student 21 about getting teams to work more effectively together. The AA follows this up with a message about the role of the facilitator in helping groups. Student 23 also follows up student 21's message by describing their problem and asking for recommendations on how they should facilitate their group. The most relevant message to this has already been posted. The AA instead retrieves a message on using Ideagenerator (a software package the students receive with the course) in a group problem solving session. The next two messages from tutor C occur on the same

horizontal (they were posted together from the offline reader) and the AA then responded to both together. The tutor's message (which was a reply to student 21) told them not to blame company culture and just get on with it. The AA responded with a message about what to do if your company does not support your studies and you are finding it hard to try out some of your ideas because people feel they are helping you with your homework. The tutor's other message (which was a reply to student 23) was about how they could help facilitate their team. The AA replied with a message about a possible structure for the second TMA. The relevance of the last two messages can be questioned to some extent. The second to last message picked up on the theme of 'my company is not helping me' without the correct context. The last message was relevant to the overall thread, which was how the student should actually go about doing their TMA, but without specific relevance to any one message.

All these messages on facilitation caused tutor C to start a separate thread, subject "Facilitation", about facilitation and some of the issues raised by the Active Archive. The message started "The Active Archive has thrown up some issues around using Facilitators - external and internal" and proceeded to give some general guidelines. The only response to this message was from another, non-electronic, tutor who elaborated on the guidelines. This is an example where the tutors did more work because of the Virtual Participant, although I would suggest that this then provided more value to the students. This was the only time in this study where there was a separate discussion directly about the content of AA messages.

A number of questions were asked of each message from the AA. Those messages which provoked the most interest and had 'common interest' questions are shown. The fourth AA message had one question which reached the status of 'common interest'. The sixth AA message had four questions, three of which reached 'common interest' level together, and the last reached this level later on. Finally the eighth AA message also had three questions which reached the 'common interest' level.

In addition to the examples of stimulated discussion and stimulated participation, and the issue of extra work for the tutors, this thread is of particular interest for two more reasons. The first is that even though there was a large number of messages from the AA this did not put the students off discussing things, and it perhaps helped the discussion for student 22. This is supported by the number of questions reaching 'common interest' level, which is an indication of the amount of additional interest that the messages from the AA generated. The second reason that this thread is of particular interest is that it highlighted a deficiency of the VP system in the second prototype. From the third message posted to this discussion the VP replied to every subsequent

message on 9 consecutive occasions. To counter this, after the sixth consecutive AA message I suppressed some of its possible matches slightly so that they would require more matching messages to reach the threshold values. However, the subsequent messages touched on so many relevant keywords that the messages were posted anyway! The messages finally stopped because none of the other threads stored by the AA were relevant. This was the first thread involving the AA during the second TMA, and by the time the ninth AA message had been posted neither of the other two threads that the AA subsequently got involved with had yet started. The problem was that with the third student message more than one thread was matched above its threshold. However the VP system only takes the top-most and posts that. Because it always maintains an overview of the whole thread this meant that additional matches were posted after every subsequent student reply. The main problem with this thread was that the subsequent student replies also added extra keywords pushing yet more threads over the threshold. Because of this some of the matches posted were not directly related to the student's message, as at each point the AA was maintaining this whole thread overview. The problem was exacerbated by the fact that the threshold values were lowered for the second study to promote more interactions from the AA.

*Problem Solving*

The final thread to consider is "Problem Solving", figure 7.5, which was started by a tutor posing the question "do we ever solve problems or do we just apply maintenance to them?". The discussion progressed on to how the course is very good at supporting the idea generation stage, but not so good at supporting implementation of the proposed solution. The first message from the AA was about what you could do if your colleagues won't help in your problem solving process. The questions in this message provoked a lot of interest from many students leading to all three of them reaching the 'common interest' level. Only one of these messages is shown because it was relevant to the later debate. The discussion continued about the various techniques and their usefulness in practical situations, and what was necessary for their successful implementation. The second message from the AA was about evaluating the techniques, which was part of the work required for the TMA. As the discussion continued further topics of company culture and how many people were involved in decision-making were raised. The third message from the AA was about how to get a more interesting job; this is a common problem that students on this course attempt to address with the various techniques covered in the course. The discussion then moved on to whether effective management is instinctive or can be learned. The AA matched a message about the I-Change technique for looking at problems from different angles. One of the 'common interest' messages stemming from the first AA message was about

'boundary analysis', and this is the only one shown in the diagram. In the second example of stimulated discussion from the second study one student responded asking to get more details about this technique. Both another student and one of the tutors replied to them explaining more about it, this being the second example of more work being done by the tutors (in this case tutor D). In this instance this student had already contributed once to the discussion conference, but only just over 3 hours previously. As this was a discussion in the second TMA over a quarter of the course had already passed.

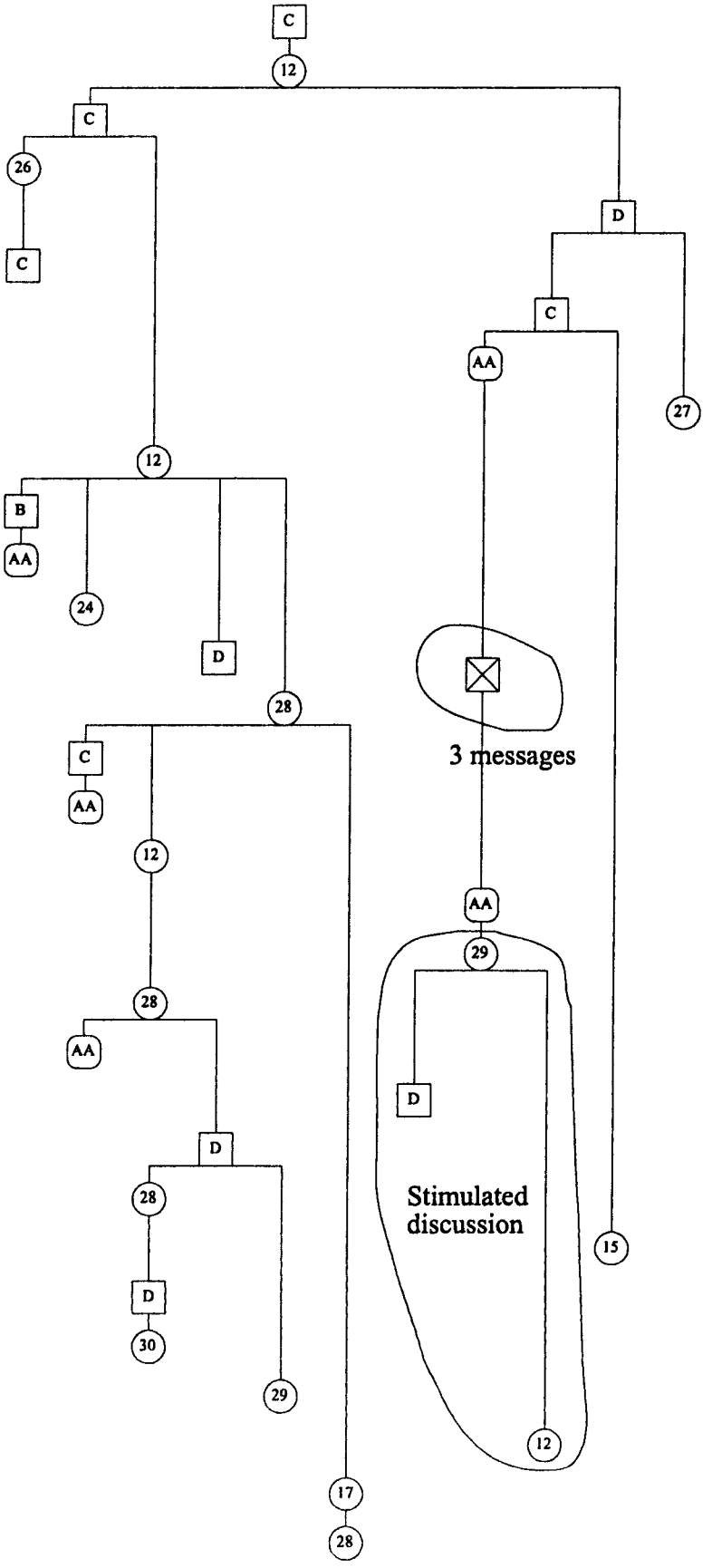


Figure 7.5: Thread diagram for "Problem Solving".

*Summary*

This section has looked in detail at the discussions the Active Archive participated in during the second study. A wide range of behaviours have been observed demonstrating the range of interactions the students experienced. The second TMA proved especially interesting as it provided two examples of discussions which were stimulated by messages from the Virtual Participant, which in both cases led to follow up by the tutors. The first example of stimulated discussion was also an example of stimulated participation, although it is impossible to tell how much the student participated either outside the public discussions or before the study began. However, I can find no trace of messages from them before this contribution within the conference archives, and as this interaction took place during the second TMA it would seem likely that they would have already contributed before this.

In this second study there has also been a case of a student posting a question publicly to the VP as their first conferencing contribution. This seems to be the only example of someone who is a lurker actually asking the VP questions. This indicates that there is indeed potential to draw people into the discussion. The public posting of questions is also interesting because, like in the first study, this means there are people who do not feel put off by doing this, although it is unclear whether in this instance it was not just an accident.

The subsequent sections in this chapter first look at the influence the Virtual Participant might have had on discussions, and then move on to consider the feedback received and the survey that was carried out.

## 7.3 The impact of the Virtual Participant on the structure of discussions

The thread diagrams were developed as a method of mapping out the structure of discussions so that the way the discussion developed over time could be visualised. If a more quantitative analysis of thread structure could be developed, based on features like amount of branching, depth, and probability of receiving a reply, then it would be possible to understand on a structural level the effect of an intervention like the Virtual Participant. There are various issues tied up in this because it is unclear whether there is such a thing as a typical discussion structure, and whether it is contextually dependent on the topic, the cultural norms for that arena, or the tool employed to support the discussions. Perhaps if such a system could be developed then a model, or a set of heuristics, of discussion patterns could be produced which a system like the VP

could use to time its interventions. I can not answer these questions here but I include some of my analysis and ideas here based on the opportunity the large amount of interactions which have taken place in this second study afford me. I have also chosen not to confront the issue of time dependence at all as there are far too many outside influences based on the course schedule to take into account.

During the first study Uncle Bulgaria had little interaction, providing only a very small data sample. Because of this it was necessary to rely on the qualitative data from the survey, which did not indicate any negative effects. This year the Active Archive's greater interaction provided more scope for a quantitative analysis.

The proposed approach for quantitative analysis of the influence of the VP on a discussion thread is to look more closely at the structure of the interactions. The representation chosen in this thesis, see figure 7.4 for a good example, is a 'tree-like' structure, in the same way a taxonomy or computer file-system is often visualised. A typical B882 First Class discussion structure would be described as deep and narrow, i.e. turn-taking discussions with little branching. The fact that these structures have low amounts of branching suggests that users may find it more difficult to follow highly branching discussions, and perhaps once the discussion is established they feel it is not their place to contribute. Depth of discussion also appears to be naturally limited. The First Class conferencing system will support depths of up to at least 128 messages - but most threads seem to die out almost as soon as they reach double digits[1]. The opportunity for late contributions is also limited as moderators move threads which appear to have reached their natural end to sub-folders in the conference. Due to a quirk of FirstClass messages which have been moved in this way can no longer be downloaded by the offline reader, so many students can no longer participate in them, and can only read them on-line.

Looking at the discussions that the VP was involved in during the second study, it turns out that they were some of the longer threads. It is unlikely that this was because the VP promoted more discussion. A more plausible explanation is that because the VP attempts to address common topics of interest, longer threads present more keywords and phrases, and an increased opportunity to match against something in the case-base. Out of the 5 longest threads in the first TMA the VP sent messages to the 1st and the

---

[1] An important factor here is that FirstClass indicates the level by showing it in the subject, i.e. "Re(10):". Most Usenet news systems do not do this and their threads often exceed a depth of 10.

3rd. Out of the 5 longest threads in the second TMA the VP sent messages to the 1st, 3rd, and 5th.

First of all, in looking at the influence of the VP on discussions I have observed that students very rarely reply publicly to any of its messages. Choosing this as the first issue to analyse let's look in detail at the number of replies received to any one message. The problem here is to select a representative sample of threads. Taking the four long threads to which the VP has contributed it is then necessary to find a comparable set of four other long threads to which it has not contributed. The sample chosen is shown in table 7.3.

| Thread Name | Depth | Number of Messages |
|---|---|---|
| Four threads for the VP, two from each of the TMA's. | | |
| Creativity or Relativity | 10 | 21 |
| Entrepreneurial, Champion and Creative leadership traits... | 10 | 33 |
| Problem Solving | 10 | 34 |
| Team Building | 7 | 40 |
| Four threads for the norm group, all from the first TMA. | | |
| Custodians of Creativity | 10 | 27 |
| Why are some people creative | 9 | 19 |
| Questioning Creativity | 5 | 21 |
| Creative Managers | 11 | 50 |

Table 7.3: Threads used as the sample for the analysis of branching.

Because the VP contributed to the longest threads in the second TMA the only equivalent sample size, in terms of thread length, was from threads in the first TMA. As thread length followed the norm presented in figure 3.2 the effect of the VP tending to be involved with all the longest threads severely limited the number of comparable threads. The assumption behind comparing between TMAs was that there was no significant change in discussion format between the TMAs.

Taking this sample we then look at the number of replies each message in every thread received. The sample group containing messages from the VP is also subdivided into those messages actually from the VP, and those from other participants. The results of this breakdown are shown in table 7.4.

| | 0 Replies | 1 Reply | 2 Replies | 3 Replies | 4 Replies | 5 Replies | Total |
|---|---|---|---|---|---|---|---|
| Non-VP | 33 | 29 | 18 | 6 | 2 | 2 | 90 |
| VP | 22 | 2 | 0 | 0 | 0 | 0 | 24 |
| VP+Non-VP | 55 | 31 | 18 | 6 | 2 | 2 | 114 |
| Norm | 47 | 45 | 13 | 6 | 3 | 1 | 115 |

Table 7.4: Number of replies to each message in the sample group.

In practice, we will concentrate only on whether a reply has been received. These summary values are shown in table 7.5.

| | No replies | Replies |
|---|---|---|
| Non-VP | 33 | 57 |
| VP | 22 | 2 |
| VP+Non-VP | 55 | 59 |
| Norm | 47 | 68 |

Table 7.5. Number of messages in each group which have or have not received any replies.

Taking the values in table 7.5 we can apply a chi-square comparison test to the data. In the case of table 7.5 this involves creating a 2 by 2 matrix with columns 'No. replies' and 'Replies' and taking the rows a pair at a time with just one degree of freedom.

| | **Table 7.5** | P-Value |
|---|---|---|
| VP+Non-VP *vs* Norm | 1.26 | >0.1 |
| Non-VP *vs* Norm | 0.37 | >0.1 |
| VP *vs* Non-VP | 22.95 | <0.001 |
| Norm *vs* VP | 20.49 | <0.001 |

Table 7.6. Results of chi-square test on the date from table 7.5.

From these results we can see that there is no significant difference between the threads in the table 7.3 groups, and especially between the Non-VP messages and the Norm group. This is a highly favourable result indicating that the influence of the VP's messages on the number of expected replies to a student's message is negligible. The other results show a highly significant difference between the VP's messages and those in the same thread, and those in a comparable norm thread. This result confirms the observation that VP messages are highly unlikely to be replied to in public discussion.

Taking the results from the chi-square test and looking at the raw data (which is legitimate in this case as the sample sizes are almost identical) we observe the following phenomena:

- VP messages get very few replies.
- VP messages tend to be leaf nodes, increasing the branching of the thread.
- Threads with the VP involved tend to have more messages with 2 replies.

This is analogous to the pruning effect in gardening where cutting a branch may make a plant sprout several more, compared to leaving it alone and having it branch rather less. So the VP's messages do not seem to disrupt the discussion, although there is a pruning effect with a slight reduction in single responses and an increase in double responses.

*Summary*

This section has looked at the effect the VP has on the structure of discussions, the conclusion being that it doesn't affect the overall structure of the discussion significantly. This conclusion has been reached by analysing the number of replies a message might receive. There are other, more general questions, relating to the structure of discussion threads and other ways of assessing the VP's impact, which are discussed in appendix 9.

## 7.4 What did the students think of the Active Archive?

The second survey, shown in appendix 10, helped to provide a good overall picture of the student's views, and provided a comparison with the first study. As with the first survey there were three sections, the first dealing with an overall view of the students' opinions on and exposure to electronic conferencing on the MBA, the second with the students' views on B882 specifically, and the third looking at their views on the Active Archive. Where possible I re-used questions from the first survey so that it was possible to make a comparison between the studies. Due to changes in the system it was not possible to do this in every case. The approach for gathering data also changed with the abandonment of the easy-to-use, MBA-student-with-little-time-friendly, tick boxes. This time the order of the scale was reversed for each successive question. This made it possible to identify questionnaires which had been filled in without actually reading the questions, and eliminate them from the study.

The sample group selected was the set of 402 students all of whom had read at least one message from the AA in the discussion conference, identified in section 7.1. This group included a wide range of those who had read messages from the VP, and asked it questions. Of this sample 123 (31%) returned questionnaires, 100 UK students and 23 from overseas. By group there were 37 identified participants, and 70 identified lurkers, the remaining 16 preferred to remain anonymous. It was necessary to eliminate 2 of the questionnaires from overseas students, and 4 from UK students. Three of these were eliminated because of inconsistent responses highlighted by the reversed scale for each question. Of the remaining three one was returned with a note informing me that they did not have enough time to complete it, and the remaining two failed to even complete the first section.

The rest of this section looks at the responses to the questionnaire. A lot of data was returned from the students; as in the first study much of this was not relevant to the assessment of the second prototype. Because of this only those results which are relevant to this study are included.

### 7.4.1 Student conferencing behaviour

A number of questions in the survey attempted to elicit some of the views and behaviours of the students relating to electronic conferencing. Both for the MBA overall, and for B882 specifically, the students were asked whether there were times when their usage of conferencing had markedly increased or decreased, to which 48% said yes (questions 3 & 10). Generally there were two events which drove these changes. Firstly, compulsory conferencing components increased usage, although the actual value of these was questioned. Secondly, as the TMA deadlines approached students either increased their usage to trawl for or discuss ideas, or decreased their usage to be able to concentrate on their other work. This shows a clear assignment-driven usage of conferencing, rather than general discussion as many tutors would wish.

The students were also asked what they regularly used FirstClass for. The most common responses to this were: conferencing, 93; to e-mail other students, 67; to e-mail their tutor, 55; for synchronous chat, 8. The number who mentioned conferencing is interestingly high when compared to the 37 who were positively identified as having interacted with the AA. However I have not been able to assess the amount of conferencing which takes place within the tutorial group conferences, and I understand from the tutors that for some groups this can be quite a high volume. It is also

interesting to note the low interest in using the synchronous chat facility. This seems to this indicate that one-to-one and conference phone calls were considered better.

Within the questionnaire there were two other measures of conferencing behaviour. The first measured students' perceived usage against their actual usage (questions 1 and 2). This can be seen in table 7.7 and seems to reflect students' unwillingness to classify themselves as heavy users even if they login once a day. It would be interesting to investigate further how these students would classify someone else as a 'heavy user'.

| | Every day | Twice a week | Once a week | Less often | Total |
|---|---|---|---|---|---|
| Heavy | 7 | 2 | 0 | 0 | 9 |
| Medium | 13 | 26 | 4 | 1 | 44 |
| Light | 0 | 21 | 24 | 19 | 64 |
| Total | 20 | 49 | 28 | 20 | 117 |

Table 7.7: Amount of use against self-classified user type.

The second measure of behaviour was a look at which areas of conferencing were perceived as important to the user (question 11). A valuable aspect of this measure is that it permits the evaluation of how useful, relative to the other conferences, the students found the Archive conferences to be. The results can be seen in table 7.8 which places the Archive in fourth place just fractionally behind the common room 'chat' conferences. This is not a bad result, and it is quite heartening to find the Archive conference valued highly by so many!

| Rank: | 1st | 2nd | 3rd | 4th | 5th | Total | Average |
|---|---|---|---|---|---|---|---|
| Coursework: | 85 | 18 | 1 | 1 | 0 | 105 | 1.22 |
| Locals: | 15 | 37 | 26 | 6 | 6 | 90 | 2.46 |
| Common: | 3 | 19 | 22 | 19 | 14 | 77 | 3.29 |
| Archive: | 2 | 19 | 22 | 20 | 14 | 77 | 3.32 |
| Computing: | 3 | 6 | 9 | 23 | 32 | 73 | 4.03 |
| Total | 108 | 99 | 80 | 69 | 66 | | |

Table 7.8: Relative importance of each conference area.

### 7.4.2 Cross-year comparison of opinion of B882 conferencing

Questions 16 to 22 were very similar to those in the first section of the first study questionnaire and focus on the students' opinions of conferencing on B882. In the first

study the students were presented with statements for which they then chose a response from a four point scale. In this study they again had a four point scale but this time the statements were rephrased as questions. (In the first study there was one extra statement about how useful the special computer tutors had been. This was left out of the second study.) The full data for these results can be found in appendix 11; here in table 7.9 you can see the average scores for each year (1 = very easy, 4 = very difficult). Taking each question pair in turn a cross-year comparison can be made to look for any changes in opinion on FirstClass. Using a 2 by 4 chi-square test for each question/statement pair it is possible to compare the answers given by the students, the results of which are in table 7.9. In this case the questions from the second study have been used.

| | Average score for first study | Average score for second study | chi-square | p value |
|---|---|---|---|---|
| How easy is it to find the right conference | 2.08 | 2.16 | 5.47 | $p > 0.10$ |
| How easy is it to find out what each conference is about | 2.25 | 2.29 | 0.65 | $p > 0.10$ |
| How valuable is the content of the facilitated discussions | 2.13 | 2.47 | 7.2 | $p < 0.10$ |
| How valuable is the content of the free discussion areas* | 2.45 | 2.78 | 7.88 | $p < 0.05$ |
| How satisfied were you with the B882 conferencing | 2.23 | 2.33 | 3.07 | $p > 0.10$ |
| How useful did you find the e-mail facility | 1.87 | 2.02 | 3.12 | $p > 0.10$ |
| How useful did you find the wider MBA conferences* | 2.13 | 2.48 | 29.49 | $p < 0.001$ |

Table 7.9: Comparison of questions on B882 conferencing across studies. Asterisks mark statements with a significant change between years.

These results are partly encouraging and partly discouraging. They are encouraging in that for many of the questions the change between the years is not significant. This is particularly encouraging for the question about satisfaction with conferencing, and the questions about how easy it was to find conferences, and find out about conferences. The overall similarity in the answers between studies seems to justify comparisons between the cohorts.

### 7.4.3 Cross-year comparison of the Virtual Participant.

In the third section of the questionnaire the students were asked about the Active Archive. These questions were more far ranging than those asked during the first study. Taking the points for development of the second system from section 5.6 this section will look for variations between the years to assess the success, or otherwise, of the changes.

*Interaction with users*

One of the key changes between the studies was the name given to the Virtual Participant. In the first study it was decided to have a name-as-metaphor approach by calling it Uncle Bulgaria. Although many students liked this name not all were so favourable. In both years the statement "The name *<virtual participant>* is a good choice" was investigated. The results of both years are shown in table 7.10, where you can see an improvement in the average score for the name 'Active Archive'. The result of a 2 by 4 chi-square test is 8.07, which is greater than the 0.05 significance level of 7.82, indicating that this is a significant improvement.

| | Strongly Agree | Agree | Disagree | Strongly Disagree | Average Score |
|---|---|---|---|---|---|
| Uncle Bulgaria | 3 | 16 | 9 | 4 | 2.44 |
| Active Archive | 4 | 50 | 14 | 1 | 2.17 |

Table 7.10: Student scores for the statement "The name *<virtual participant>* is a good choice" from each study.

One of the reasons behind the change in name was to address a comment from the first study that *"It needs an unusual name so it is clear that it is not a student or a tutor"* [comment 6, appendix 5]. In making the change it was intended to make the link between the 'Active Archive' and the 'Archive Conference' more explicit. In section three of the questionnaire the students were also asked about how much use they had made of the archived conferences. There seemed to be some confusion in answering this question as to whether it was about the 'Active Archive' or not. Again, this highlights the importance of selecting the names used very carefully.

Other comments made by some students also raised certain worries[2]. It had been made clear to me by the water rats that one of their concerns was about students placing undue weight on the messages from the AA. Student comments show this could indeed be a problem. Two students said; *"assumed relevant ones would be posted to discussion conferences"* [24-14], and *"used past tutor comments and read all questions"* [24-11]. A later question in section three of the questionnaire asks students if they felt they learnt anything from the AA messages. One student responded saying *"Very relevant messages - seriously impressive database management.* [but when it came to learning from it] *not really - the cleverness of its being there means it has undue weight attached to it, whereas in fact its content is only the same as that of the current discussions"* [34-11]. This serves to reiterate the need to make the role of a system like the Virtual Participant, and the background to its information, very clear.

One other aspect of student - VP interaction considered last year was whether they preferred a public or private method for asking questions. The results of the survey last year when asked to rate the statement "I would prefer to ask UB questions directly by private e-mail, rather than through the conference" showed no significant preference (section 5.1). This year the students were asked to rate the statement "I would rather the Active Archive posted messages directly to those participating in the discussion and not to the conference". The answers to this question can be seen in table 7.11, and a 1 by 2 chi-square results in 9.61, which shows significant disagreement with this statement ($p < 0.01$). This indicates the value attached by the students to their seeing public messages from the Virtual Participant, even if they were not willing to ask questions in public.

| | Strongly Agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|
| I would rather the Active Archive posted messages directly to those participating in the discussion and not to the conference | 0 | 20 | 38 | 7 |

Table 7.11: Answers to question 40 in the second questionnaire.

*Format of the cases*

In this study, as in the first, we did not explicitly ask for comments on the presentation of the messages from the Virtual Participant. Having addressed the technical problems

[2]All of the comments referenced in this chapter, unless otherwise noted, can be found in Appendix 11.

from the first study we received no complaints. This is classed as a 'consumer expected feature' which the users would expect as a minimum from the system and complain bitterly if it was not there.

As part of the format changes we also exposed the case-base in the form of the 'Archive' conferences. Although there was obviously much interest in, and use of, these conferences we did receive some complaints that the structure made it difficult to find things. During both studies comments have been made by many students that a search engine of some kind would be useful. Others asked again for a system which would summarise discussions in the conferences, highlighting only the important bits in the way a recommender system would [23-11, 25-13, 29-4, 25-11, 24-16]. Although I can see the benefits students might perceive in these additions they do not necessarily increase the amount or quality of conferencing. Experiences with recommender systems (Resnick & Varian, 1997) would suggest that such additional functionality might even reduce and devalue discussions in the conferences.

*Retrieval*

With the second prototype it was the intention to build delays into the messages sent by the VP, to allow the tutors to monitor its current state, and to incorporate an approval mechanism. The ability to delay messages from the VP was implemented, but disabled because of changes in the conferencing structure. Of the other two points, both were technically possible, and seemed necessary to win the approval of the tutors. However when the study started it was clear that the tutors were happy with the system anyway, and no longer wanted that level of control.

To evaluate the relevance of the retrieval process the students were asked, in both years, how relevant the messages from the Virtual Participant were to the current discussion. The answers to these two questions can be found in table 7.12. A 2 by 4 chi-square test on these results shows that there is no significant difference between them.

| | Strongly Agree | Agree | Disagree | Strongly Disagree | Average Score |
|---|---|---|---|---|---|
| Uncle Bulgaria | 4 | 29 | 3 | 0 | 1.97 |
| Active Archive | 1 | 44 | 9 | 0 | 2.15 |

Table 7.12: Student scores for the statements "The AA messages were relevant to the current debate" and "Uncle Bulgaria's messages were relevant to the discussion".

The other aspect of retrieval is the impact messages from the Virtual Participant might have had on others in the conferences. In both studies we asked the students whether messages from the Virtual Participant put them off participating or sending their own messages. Table 7.13 shows three questions, one from the first study and two from the second.

| | Strongly Agree | Agree | Disagree | Strongly Disagree | Average Score |
|---|---|---|---|---|---|
| 1) UB's messages put me off sending my own | 1 | 7 | 21 | 5 | 2.88 |
| 2) The AA put me off participating in the conference | 1 | 8 | 38 | 17 | 3.11 |
| 3) The AA reduced the amount of discussion | 0 | 8 | 40 | 5 | 2.94 |

Table 7.13: Student scores for questions about the impact of messages from the Virtual Participant.

Comparing 1 and 2 in table 7.13 gives no significant difference ($p > 0.1$). This indicates that there has not been a significant change in the number of students feeling put off by the Virtual Participant, however the trend is towards less. Statement 3 is included in the table as it is related to numbers 1 and 2. A 2 by 2 chi-square comparing agree and disagree between 1 and 3 shows no significant difference ($p > 0.1$). A 1 by 2 chi-square on number 3 indicates there is significant disagreement with the statement "the AA reduced the amount of discussion" ($p < 0.001$).

These results show that there has been no significant change in the proportion of students who feel that the Virtual Participant has put them off participating. However the numbers who state that they were put off participating are quite small. These indicate an important group whose numbers should be minimised for systems such as the VP.

Although this group of students is small, for wide acceptance of a system such as the Virtual Participant it is necessary to address the problems of even these few students. Unfortunately the results from the questionnaire are not detailed enough to determine precisely why these students felt put off from participating. However, it is worth pointing out that even these students seemed to feel that the VP had a role in the course. Looking at their agreement to *"The AA should continue to be used on this course"*, shown in table 7.14 even they still seem to feel it should continue to be used.

| | Strongly Agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|
| The AA should continue to be used on this course | 0 | 6 | 2 | 1 |

Table 7.14: Answers to the statement *"The AA should continue to be used on this course"* for the students who felt the AA put them off participating.

This group is not big enough to do a meaningful statistical evaluation, but it would seem to indicate that even they feel that the system should continue to be used. The student who strongly disagreed with the statement that the AA should continue to be used was the same one who strongly agreed that they were put off by the AA. A further study into the opinions of these very few students would help to discover why they felt the way they did, and help improve the VP to address these problems.

The question about whether the VP should continue to be used was asked in both years. The results can be found in table 7.15. Although the average score indicates stronger agreement in the second study a 2 by 2 chi-square test comparing agreement with disagreement shows the difference is not significant. However, within both years comparing agreement with disagreement (a 1 by 2 chi-square in both cases) shows significant agreement ($p < 0.001$).

| | Strongly Agree | Agree | Disagree | Strongly Disagree | Average Score | chi-square |
|---|---|---|---|---|---|---|
| Uncle Bulgaria | 6 | 18 | 4 | 2 | 2.07 | $p < 0.001$ |
| Active Archive | 9 | 55 | 6 | 1 | 1.98 | $p < 0.001$ |

Table 7.15: Cross-year comparison of student responses to the statement "The *<virtual participant>* should continue to be used on this course".

This concludes the subsection on retrieval, which has also looked at student interaction with the Active Archive. There have been no very significant improvements over the first year. However, given the increased level of participation of the Active Archive in comparison to Uncle Bulgaria it is important to see that there has not been a significant increase in the numbers of students feeling put off. It is also good to see an increase in the average agreement with continued use of the Virtual Participant, even if this is not significantly greater than in the first study.

*Maintenance*

One aspect of the Virtual Participant which is vital to the continued use of systems such as this is maintenance. The knowledge base needs to be maintained, something which can not be done in a fully automated fashion. It seems likely that the best method of doing this is to seek feedback from the users of the system, to determine the value of the existing content. However, over and above this, new content can be identified from threads of discussion which do not match with the contents of the current knowledge base.

In the second study it was intended to use the revision process before the exams to provide feedback about identified key points in the course. Unfortunately, given the tutors' lack of time and workload this was not possible. Of course, the original method for producing the case-base could be reapplied to the conferencing discussions, enabling new threads to be added and others updated. However this is not an ideal solution as the responsibility and work still falls to one person.

The main change made to support the maintenance process in the second version of the VP was to give students the ability to provide feedback. For every message the students could rank two statements, and then give any other comments:

- Relevance to Discussion: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ].
- Useful for You: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ].

The first was intended to gauge how accurate the VP had been in matching this message. The second was to gauge how useful the student themselves found it. Usually those who chose to give feedback gave good feedback. However when giving feedback students would often give the relevance to the discussion a high mark, even when it obviously wasn't particularly relevant.

In the survey students were asked "How often did you give feedback on messages from the Active Archive?", and also asked "If you did not give feedback, why was this?". The results for these are given in table 7.16.

| | Always | Often | Sometimes | Never | Average |
|---|---|---|---|---|---|
| How often? | 1 | 2 | 17 | 60 | 3.70 |
| | Always relevant | Too much effort | No benefit to me | other | |
| If not, why? | 10 | 17 | 18 | 4 | |

Table 7.16: Student responses to questions 30 and 31 of the second questionnaire.

The responses shown in table 7.16 show that three quarters of the students responding never gave any feedback to the Active Archive messages. There were ten [31-1 through 31-10] comments made by students as to their reasons for not giving feedback. Four of these said that they did not know how, or had tried and failed. Again I think this was perhaps due to students using old versions of the offline reader which did not provide the right functionality. One student said that as they were paying to use FirstClass they were unwilling to pay to give feedback. This is a long term problem with the use of FirstClass and part of the reason for moving towards a web interface for students, so that they can access it more easily from work. Two students made comments about interacting with the Active Archive, one stated that it was *"too impersonal"* [31-10] and the other that *"talking to an automation, that's really sad"* [31-5]. This serves to highlight some students' perceptions of interacting with a system such as the Virtual Participant, and it is something that needs to be studied and understood better to increase the amount of feedback. Students need to see some advantage or reward to giving feedback. Finally, an important comment from one student was that *"no feedback re. receipt of messages left me doubting that 'feedback' was actually used"* [31-7]. This is a very important point to consider in looking at the motivation of students to provide feedback. In no way was their contribution acknowledged, and in this case it may be that if it had been, this student would have continued to give feedback.

Addressing this question of how to keep knowledge in a community up to date is a key current research topic. Questions about motivation, reward structure and the like have yet to be addressed. Within this research I have not been able to properly study this problem, and it seems to be especially difficult in this context as the student community has no commitment to those in future years, and the tutor community does not see an immediate benefit. The next chapter looks in more detail at the lessons learned from

this work and provides some practical requirements for systems like this as pointers for further research.

*Tutor's Assistant*

One of the long standing motivations behind the Virtual Participant was for it to be an assistant for the tutors. With the Active Archive implementation much of the proposed functionality is available. However the major stumbling block remains the acceptance of these ideas by the tutors. The second study has shown that as they became more experienced with the system their request to be able to control it and be aware of its operation, although achieved, proved less necessary as their trust increased. It is perhaps unfortunate that this research can not be continued with another group of students, as the increasing acceptance of the system shown by the tutors would indicate that some of the additional functionality offered by an assistant might begin to prove useful.

*Summary*

This section has taken the results of the second study and compared them with the first to assess the changes in design. The results have all been positive, showing either a significant increase in acceptance since last year (e.g. for the name) or no significant change in opinion in the face of increased involvement on the part of the VP (e.g. in the numbers of students feeling put off). This indicates a successful second study where improvements have been made with significant benefits and the range of interaction of the system increased without detrimental effects.

### 7.4.4 Student perceptions of the Active Archive

The last questions in the survey were intended to gauge students' perceptions of the Active Archive. This was approached by asking them to write down a few words to describe their feelings about certain aspects of it. A list of words was suggested, but they were not constrained to use them.

Focusing again on the name, one question asked the students to pick words to describe the name 'Active Archive'. Figure 7.6 show the responses given by those students, the top 6 of which are all positive. To enhance readability of the figure all those words chosen once have been omitted, and can be found in appendix 12. Following on the next question asked the student to write down what they felt when they saw a message from the Active Archive. Figure 7.7 shows the responses given.

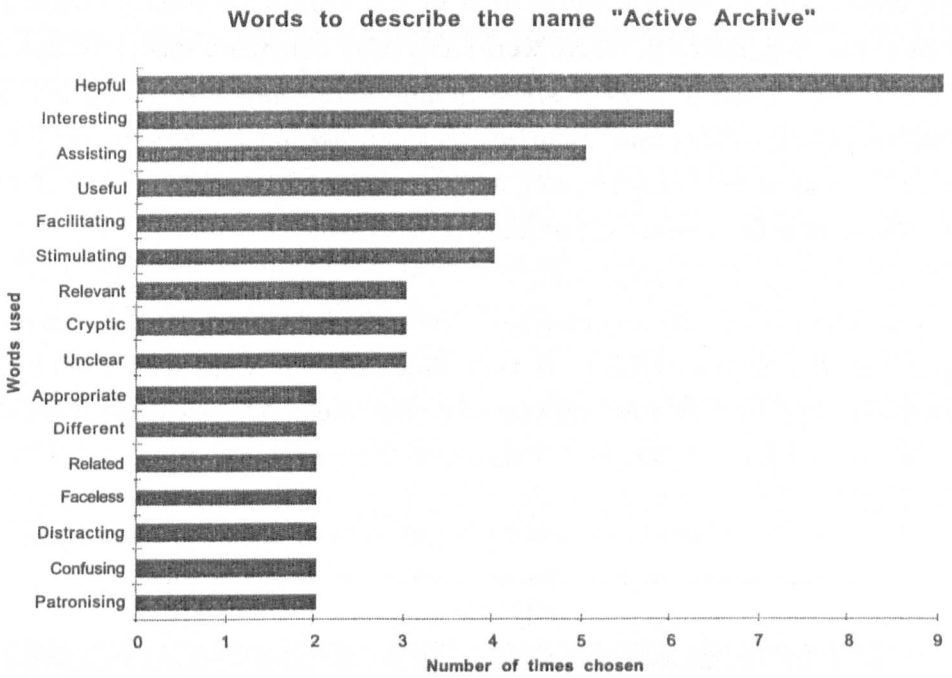


Figure 7.6: Words selected by students to describe the name "Active Archive".

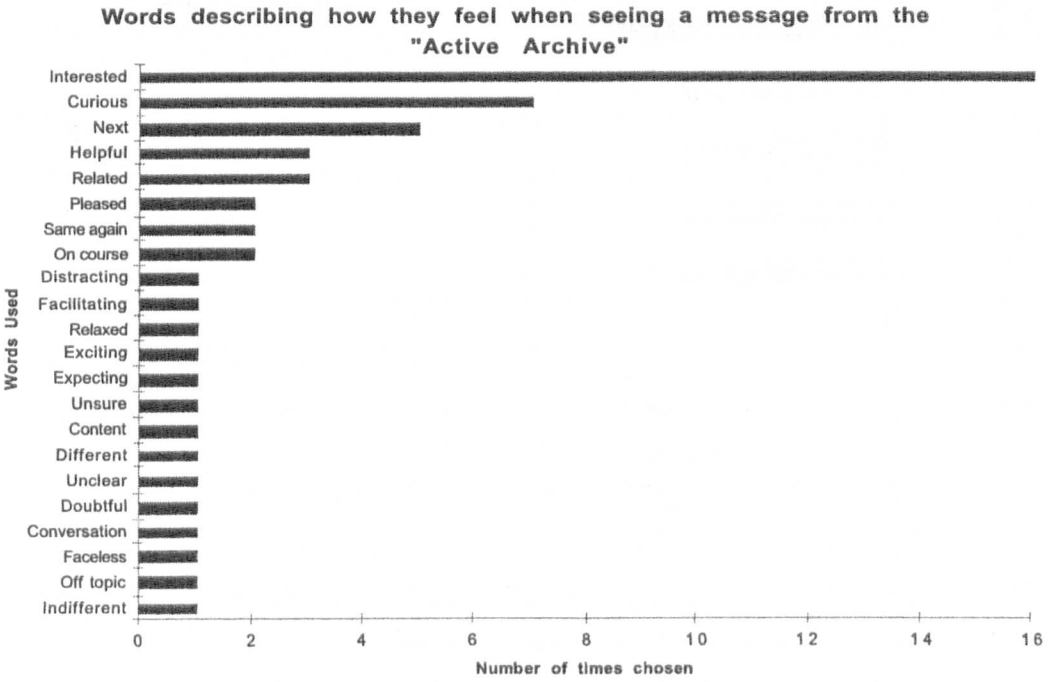


Figure 7.7: Words selected by students to describe how they feel when they see a message from the "Active Archive".

In these tables words with the same stem or meaning have been grouped (i.e. interested, interesting, interest). There were some other comments made, for example one student said *"Usually read it since it has more chance of being a useful contribution"* [46-2], another said *"Positive - look at it like any other message"* [46-1]. The second comment here is heartening as it has always been the intention for the Virtual Participant to be considered 'just another participant'.

The next question asked the students for a few words to describe the content of messages from the Active Archive. A very wide range of terms were used by the students including 27 which were chosen only once each. Those chosen more than once are shown in figure 7.8, the rest can be found in appendix 12.

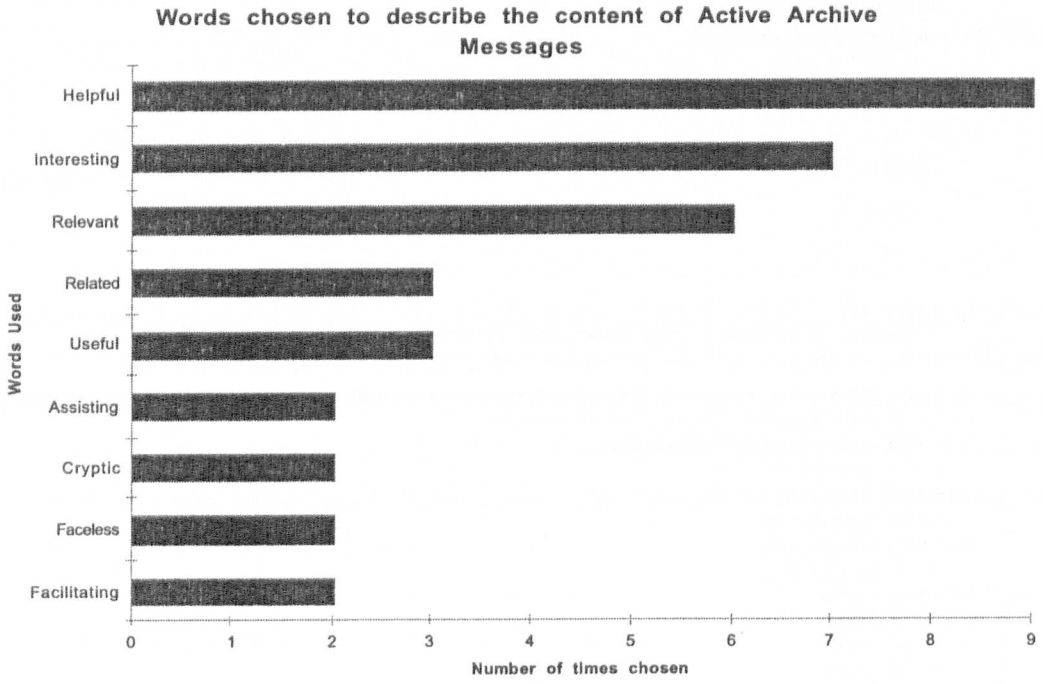


Figure 7.8: Words selected by students to describe the content of messages from the "Active Archive".

The following question looked for the students' overall opinion of the Active Archive. In response to this question a number of students made statements rather than just picking words. The words they chose are shown in Figure 7.9. The statements included two from students who said that it *"should be used more often"* [48-6], one who said that it was also *"distracting but should be used more often"* [48-4]. A slightly puzzling comment from one student said that *"good idea - each message should be summarised at start with short sentence of contents"* [48-1]. As each message had a headline phrase and a scene-setting sentence at the start it appears there is always a

demand for an *executive* executive summary. This view was upheld by another student who said *"just key infor would be better - delete superfluous stuff"* [48-7]. This is a common theme amongst MBA students, and perhaps managers in general, where they want the maximum amount of information in the minimal content. However, providing students with the summaries and search engines they ask for would, I believe, reduce the value of conferencing and probably also reduce their learning from the course. Other positive comments students made included *"informs current discussion"* [48-8], *"A good innovation with promise for other courses"* [48-3], and "*a possibly powerful tool"* [48-5]. Other students saw room for improvement, commenting: *"could be better - easier to use - more instructive"* [48-2], and that there was an *"opportunity for improvement ;-)"* [48-10]. Finally one student questioned *"clever, but over the top. Nice to know it can be done, but so what?"* [48-9]. Perhaps this best shows the need to continue to prove the worth of the approach to some students.

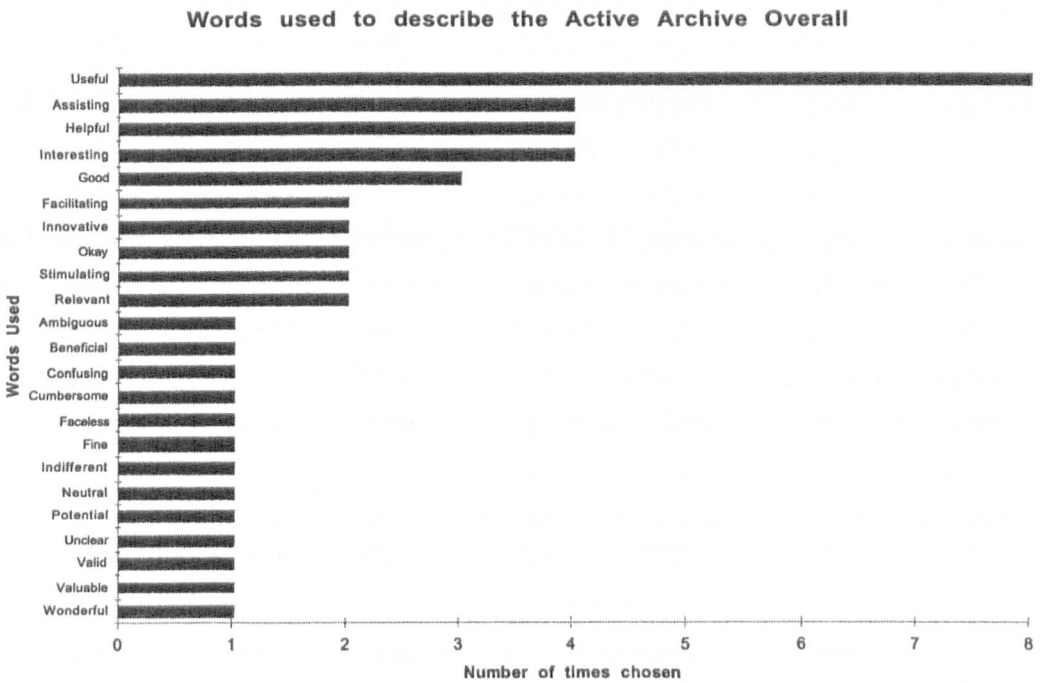


Figure 7.9: Words selected by students to describe their overall opinion of the "Active Archive".

Finally, the students were given the opportunity to make any other comments they wanted about the Active Archive. One obvious point arising from these comments was a lack of understanding on the part of some students what it was the Active Archive actually did. Students commented: *"Very useful concept but would be better explained - i.e. wasn't sure what it was for a while then got the hang of it."* [50-2], and *"As an active user of FC I have more or less ignored Active Archive - and still do not know*

*what it is trying to achieve"* [50-3]. This was despite messages to the discussion conferences from tutors informing them about it, information in the very top level B882 conference about it, and simply clicking on its name would use the *resumé* function built into FirstClass to give information about the Active Archive.

One student felt that it had helped them, providing information that they would not have found, as they would not have used the archive. This student said *"Generally, I think it is good and an asset to my learning + development. It is a massive time saver - I would not look at the archive at all, even if it were not there, and might impoverish my learning as a result."* [50-4]. This is an important statement because it shows that this student clearly feels they would not have benefited as much from conferencing without the Active Archive. I do not think that it is too presumptuous to assume that there are likely to have been other students who had also experienced this. However to maximise the overall benefit it is necessary to ensure that other students, who do not benefit in the same way, are not disadvantaged by the Active Archive.

The last two significant comments from students highlight an opinion on how it should be used, and a perception of its function. One student said that *"the active archive should probably only be used to stimulate discussion when or if it is flagging"* [50-5], which is certainly a role it is trying to fill. Using the proposed delaying function would help to use the Active Archive to bridge gaps in the conversation. The idea of it as a tutor's assistant, actually starting discussions, has not yet been tried out. From the comment given this student would probably support that approach, but it is not clear whether other students would, or (especially) the tutors themselves. This is certainly another aspect which it would be interesting to explore in a future study. Another student got quite irate about even being asked to pick some words to describe their feelings and said *"Its a tool! A dictionary is a tool! I have no interest in what you call either of them. Neither do I find either patronising, stimulating etc."* [50-1]. However in the first study, when it was called Uncle Bulgaria, there was one student who thought it was another student posting messages and hiding who they actually were. Given the limited scope for cues in text-based systems, people often interact with automated characters assuming they are human, or perhaps not realising that they are not. This has been studied quite extensively by Sherry Turkle (Turkle, 1997) and this evidence with the range of response received by the Virtual Participant in its various guises, indicates that perhaps we need to make it yet more explicit that it is a mechanomorphic entity. The next chapter will discuss this issue further.

*Summary*

Section 7.4 has taken a detailed look at what the students thought of the Active Archive and compared these opinions with those from the previous year. This section has evaluated the students' views in a different way to before, showing a common choice of positive terms by many students. There was no detailed evaluation of what the *tutors* thought of the Active Archive, and in many cases I think they simply ignored it. However it was the same group of tutors in both years and the next section briefly discusses informal interviews with two of them.

## 7.5 Interviews with tutors

I conducted informal interviews with two of the electronic conferencing tutors. Tutor 1 was chosen as they had always been willing to give their opinion, were a self-proclaimed champion of the 'novice user', and were always concerned about the effect the VP might have on the students. The second tutor was chosen as they were responsible for the management and maintenance of the FC conferences for the B882 tutors. Full notes from the interviews are available in appendix 13.

To summarise the important points from these interviews:

- The tutors were generally happy with the concept
- The implementation needs some tweaking - specifically a better set of matching heuristics, or a language understanding system
- The knowledge base is dated and needs to be kept up to date
- In this current form it does not necessarily reduce work load although the additional materials may improve the teaching
- The solution to taking it forward is getting the co-operation of the tutors

In brief, once the bugs were worked out, they liked the idea of the VP, or perhaps something like it, and see the way forward for it as through more tutor involvement.

## 7.6 Conclusions

This chapter provided a comprehensive account of the second study of the Virtual Participant. With the Active Archive the lessons learned from Uncle Bulgaria have

helped to create a system which had much more interaction with students, without an increase in negative opinions. The students gave the name change a significant thumbs up, although it was perhaps still not yet ideal. Despite increased participation on the part of the Active Archive, and a much wider sample of students, it was not seen more negatively than in the previous year on any of the measures tested.

The creation of 'Archive' conferences appears to have been popular with students with many students making use of them. However, one slightly worrying comment from a student [24-14] indicated that they believed the archive conferences to be of little value as any important messages would be posted by the 'Active Archive'. This comment, and similar points ([34-11] for example) made by other students, indicate that the profile and explanation of the purpose of the system needs to be raised even further. However, there is a risk of making some students sick of the sight of information about the Active Archive.

When it came to retrieval there were no differences between the years in the algorithm and patterns used, but the threshold levels were lowered in the second study. In response we saw a slight shifting of the average response with students slightly less likely to feel 'put off' and 'were the AA messages relevant' in comparison to the previous year. However none of these changes were significant, even in the students' opinions of whether it should continue to be used (which they agreed with slightly more than before).

It is clear that the problem of maintenance has not been fully solved. Without direct help from the tutors, by incorporating it into the study process, and identifying the key motivating factors for the students, then this will continue to prove difficult. The feedback system incorporated into this year's study has not been successful. There was no motivation for students to make use of it. Those who did use it did not receive any acknowledgements, and perhaps assumed that the feedback was being ignored. From studying the feedback messages that were received, they seem to present a distorted view based on the opinions of a few students who tended to overestimate the relevance, and therefore may be of little practical use anyway. At worst this distortion could lead to changes to the system being based on a biased sample.

In the second survey the students' perceptions of the Active Archive were also sampled by getting them to write down words in response to statements. Although it is difficult to make an objective analysis of this kind of feedback, it seems that positive terms like 'useful', 'helpful', 'interesting', and 'assisting' have occurred relatively more frequently than negative terms in the responses. This is especially important in the light

of the increased contributions of the VP in this study and the fact that it stimulated extra discussion.

This chapter began with a look at the question-asking behaviour of students, the structure of threads in the second year, and investigated the impact of the Virtual Participant on discussions. Although overall there was no quantitative evidence for increases in participation at the gross level there are specific examples where the discussion and participation was stimulated (although these cases did lead to extra work for the tutors, something the VP was supposed to avoid). These examples provide some good evidence that this system does have the potential to meet its goals.

Finally, when looking at the impact of the Virtual Participant on discussions and thread structure there is a very clear tendency of students not to respond to its messages in public. Only 2 out of 42 AA messages were responded to in public, where every AA message which had questions which could be asked had those questions asked in private. However there was no sign that, even when many messages were posted, there was any disruption to the discussion. Students' feedback from the questionnaire also supports this view that the AA did not unduly influence discussion.

The next chapter brings together the results of both studies to define the development guidelines which have emerged. It will draw together some of the practical requirements necessary for developing other systems looking to support users with targeted knowledge. Finally, ending the thesis, it looks at the conclusions reached and provides pointers for further study.

# Chapter 8: Conclusions: Supporting learning with collaborative story-telling

## 8.0 Prologue

This thesis started by looking at the concepts of learning cycles and tertiary courseware and their implementations with multiple users in computer-supported learning environments. Starting from thoughts about learning conversations, supporting dialogue, 'virtual tutors' and 'virtual teaching assistants' this research has progressed through an initial proof of concept, and evaluation of that first prototype, a second improved demonstration system, a re-evaluation and comparison of this second prototype, leading finally to key lessons learned from the development. It has not yet been possible to achieve a highly successful solution, dramatically improving the learning experience for all students, although much of the material needed for this now seems to be within reach. Even so, this thesis has, I believe, taken several significant steps along the path to better approaches for computer supported collaborative learning.

This chapter is intended to provide a final review of the work presented in this thesis. To begin, the first section will highlight the key learning points to take away. Following that are three sections looking in turn at requirements for organisational memories, usability problems, and agent based systems to support learning. Moving into the closing parts of this chapter I first review what has been learnt against the goals set out in chapter 1. Almost last, but certainly not least, I will discuss the possible future development of this work. Finally, an epilogue will bring this thesis to a close.

## 8.1 Lessons learned in the design of the Virtual Participant

From the studies four key issues have emerged which are relevant to the design of systems like the Virtual Participant. The evidence for these guidelines is highlighted here, exemplifying why such issues must be considered. Using these lessons as a starting point I will then draw parallels between this work and recent guidelines for the design of organisational memories (Abecker, Bernardi, Hinkelmann, Kühn & Sintek, 1998). The chapter will then bring in aspects of recent work from (Masterton & Watt, 1999) on social roles in knowledge management systems, and compare it with the roles of pedagogical agent systems discussed earlier in chapter 2.

This section presents an overview of these four key design issues or 'lessons' which emerged as our experiment and evaluation progressed. These lessons provide four fundamental dimensions which have proved important in the social acceptance and usability of the Virtual Participant. These four design issues are:

a) anthropomorphism versus mechanomorphism

b) private versus public

c) closed versus open

d) fixed versus extensible

These were first presented in (Masterton, 1997). Each of these sections presents the aspects of the Virtual Participant that have changed and the situations which drove those changes, drawing parallels with other systems. I refer to some specific examples with Answer Garden (Ackerman, 1994a) and the Knowledge Sharing Environment (Davies, Stewart & Weeks, 1998), both described more fully in chapter 2.

**Anthropomorphism versus mechanomorphism**

A common issue for Artificial Intelligence programs that interact with humans, particularly of the agent genre, is that they need to present themselves in a way that reflects their ability. Where there is a conflict between the ability of the system and the user's perception of that ability a breakdown occurs. The user may then fail to exploit the system's full potential or become frustrated with its shortcomings. The Virtual Participant was initially anthropomorphised, as 'Uncle Bulgaria', with the hope of presenting it as something friendly and useful, but not as fully human or having human capabilities. Two other anthropomorphised systems can be seen in NewT and Maxims (Maes, 1994) which presented visual representations of the agents' internal states. There was no evaluation of how helpful these representations were to the users. Koda (1996) found that there were large differences in users' opinions on "agents with faces". Other researchers have also found it necessary to alter the presentation of systems to improve their acceptance (Watt, 1996).

The MBA course in Creative Management, which the VP was tested on, made great use of metaphor. When faced with the task of selecting a First Class user name for the first study I originally chose 'Uncle Bulgaria', a cartoon character from a popular British children's television programme (The Wombles); a senior and knowledgeable

figure amongst a group who recycle things. Although the majority of students were neutral to the name, and some really loved it, others had strong negative feelings. As one of the returned questionnaire comments tactfully put it "the name had an adverse affect on what the objectives of UB were". When the students were asked to rate whether the name was a good choice there was some negative feeling, see section 5.1. Although this was not statistically significant, it is an important issue for acceptance, especially when the name provokes highly negative responses in some people. Given the feedback received and the specific dislike expressed it seems that this is due not the contents of the messages, but to the 'avuncular' presentation of the name.

To address this problem, in the second study the system was 'mechanomorphosised' with the name the 'Active Archive', as discussed in section 5.6. Unlike the first study this provoked no strong negative reactions, and a cross year comparison showed a significant increase in user acceptance, see section 7.4.3. In this case no attempt was made to make the AA appear as anything more than an automated search engine. When considering this it is also worth considering anecdotal evidence which suggests that younger students and children are more comfortable with anthropomorphic representations than older adults. It would certainly be worth more research in this area, especially as such representations are often used in computer software aimed at younger children.

**Private versus public**

The original concept for the Virtual Participant was that it would function like any other user, with certain caveats: it only had access to certain conferences and it could only hold public conversations. The interaction model used followed this sequence: (a) the Virtual Participant tracked student discussions attempting to match them to its database, (b) when a match was found the root message was posted, (c) students could find out more information by posting pre-defined questions to the conference, (d) the system then posted the follow-up information to the conference.

In the first prototype, and in contrast to Contact Finder (Krulwich & Burkey, 1996), all interactions with the students were public. It was felt that by keeping these interactions public we would encourage group discussion, helping students to learn vicariously, without being seen to be unfairly benefiting a small group of students.

Although in our first study over half the students said that they were comfortable with asking questions publicly, hardly any questions were asked in practice. Additionally, despite specific instructions and the fact that this was actually harder, a number of

students posted the questions by private e-mail. This was altered with the second prototype and although the Virtual Participant still accepted public questions the main method was by private e-mail. This had the potential to deprive students of the chance to learn vicariously from VP messages, so to address this I introduced the concept of 'common interest', where any question asked more than five times was posted to the conference. As can be clearly seen in chapter 7 the second study saw many more questions asked, both in number and in range, but in almost all cases these were asked by personal e-mail. The results in section 7.4.3 also indicate the value gained by having the messages from the VP publicly on view, even if the students are unwilling to ask questions that way.

In comparison to this KSE (a collaborative filtering system reviewed in section 2.3) links information to the user who submitted it. In this way it is possible to track who has added what to the database. Although this aids in finding out who added a piece of information to the database (which is important in Knowledge Management), it discourages users from adding information, especially if they do not wish to be credited with it. Answer Garden makes all question-asking of experts anonymous, to deal with possible status problems for novices bothering experts. An interesting caveat here is that the different modes of communication mean that some users are happy to ask experts questions by e-mail, but would not by telephone, or face to face. Others have tried to address this problem; for example Yenta, a matchmaking system, is specifically designed to enable users to find people with similar interests while protecting their privacy (Foner, 1996). Finally (Resnick & Varian, 1997), in their taxonomy of 'recommender' systems, compare the public, anonymous, and pseudonymous nature of these systems.

The approach finally embedded in Virtual Participant attempted to balance privacy with the wider needs of other users' learning. By allowing both private and public (even though they were rarely used) questions, this allowed the user to decide whether they wanted to be associated with the questions or answers. With the 'common interest' concept there was an arbitrary measure of how much use a certain response would be to the whole user population. This served the role of 'recommending' that information to the rest of the user group, while protecting the privacy of those involved.

**Closed versus open**

In the first study the students had no direct access to the Virtual Participant's knowledge base. When we asked the students if they wanted access to everything the

response was a significant yes (section 5.1). Therefore for the second study it was decided to allow access to everything in the knowledge base. This access took the form of archive conferences for each TMA containing a number of threads on different topics, the subject reflecting the contents. Each thread corresponded to a different topic within the knowledge base.

In the second study the students were asked to give feedback both on the relevance of the message to the current discussion and on the usefulness of the message to them. Work conducted in (Resnick & Varian, 1997) shows that there is a range of user types including some (a few) who are happy to give feedback with little or no prompting. Unfortunately, in the Virtual Participant, because there was no clear benefit to giving feedback, it was only given rarely in practice. When asked, the students said that they usually felt giving feedback was too much effort or of no benefit to them. One clear point of failure here was one student who felt that because the feedback was not acknowledged it was being ignored, section 7.4.3. Clearly this is an important point to consider: for users who are happy to give feedback, it is necessary to acknowledge it!

The real benefit of the Virtual Participant comes from contextualising the contents of the archive to the current discussion. Making the entire database available has been driven by the students' demand to have access to everything. By tracking users browsing the discussion conference and the archives it was clear that by making this knowledge more accessible more students accessed the data. It is not possible to say whether they made use of it! Other systems also have a clear trade-off between presenting identified information and allowing the users to browse everything. In Answer Garden the user answered questions to navigate to the appropriate part of the database where they were presented with an 'answer'. If this answer did not help them, the next stage was to ask an expert. However it may be that the user did not get to the right part of the database, but were in the right neighbourhood. Here an option to, for example, 'search for answers related to this one' would prove useful. Both the Virtual Participant and Answer Garden databases had initially been generated by experts, perhaps with some automated assistance, and with an application in mind. This has made them open to some extent, but adding new information can then become a problem — we will look this at in the next section. In the case of KSE everything is open and all information is available to all users, but this is at the expense of the structure of Answer Garden and the Virtual Participant.

The Virtual Participant had the advantage that its database was small enough to permit serendipitous browsing. The expert-imposed structure served to provide a

coherent arrangement of the knowledge in an accessible form. With that, the VP approach maintained the advantage of contextualising the knowledge in the database into discussions at the time of need. This removed any burden on the students to conduct searches. Unfortunately there was still a failure to get adequate feedback, so the problem of motivating users to contribute to and maintain the knowledge base remains to be solved.

**Fixed versus extensible**

One problem with the Virtual Participant was that its knowledge base was, with some automated assistance, hand crafted. To be extensible to *other* domains the knowledge base generation process must be simplified and made as automatic as possible. To be extensible *within* a domain the database must be able to be kept up to date. Answer Garden had this same problem of using an initial database hand crafted by experts. To the users of these systems the databases appear fixed, although the experts (or in the VP's case the tutors) can extend them.

Creating a knowledge-base and extending a knowledge-base are not necessarily different things. It may be possible to 'pull oneself up by one's bootstraps' so to speak. Seeding a database and providing an initial structure is, with the help of an expert in the field, quite straightforward. What needs to be addressed is who benefits from the production and maintenance of such a knowledge base? Those who benefit directly should be the tutors, or experts, by reduced demands on their time, but they may not be motivated enough, or have enough time to start the process. Students, on the other hand, benefit only from the product and have no reason or motivation to maintain it. One approach to providing them with this motivation is to make some feedback to the system by the students compulsory or assessed[1].

With the Virtual Participant, if the motivation problems can be addressed, there are several ways the updating process can be supported. As the contents of the knowledge base originally came from students' discussions each year presents fresh material. By matching discussions to threads in the knowledge base the VP is already providing a classification mechanism for knowledge. Those discussions which are not matched can then be classified and added to the knowledge base. Outdated

[1] This approach was adopted by the Vicarious Learner (Lee et al., 1997) but I am critical of this option, see section 2.4 on Tertiary Courseware.

material can be identified by its failure to match with current discussions, and from negative student feedback. However, in all cases it is still necessary for an expert to have the final say.

Like Answer Garden, the way forward for the Virtual Participant is to find ways of motivating some section of the user population to assist with the maintenance of the knowledge base, encouraging a knowledge-sharing culture. With the Virtual Participant it was proposed to do this by getting the students to reflect on what they have learned, and then to add these points to the database. Looking for evidence of the effectiveness of this approach is a topic for future work. With KSE, as with other recommender systems (Resnick & Varian, 1997), there are likely to be different groups of contributors to the database, so there is a risk of a drop in usage from high contributors if they do not feel they are benefiting. However, by encouraging the system to be used within groups of users who already collaborate this risk should be reduced.

An extensible, maintainable, system is always preferable to one with a fixed set of knowledge which may become obsolete over time. This was always a fault of traditional rule-based expert systems, where changing rules was difficult and could expose hidden dependencies. The Virtual Participant is not yet an ideal solution to this problem, but it can in principle provide a framework which will support maintenance by identifying updated knowledge and gaps in knowledge. However, until a method of motivating knowledge maintenance is found then it will always be considered a chore.

**Summary**

Here we have seen four key lessons learned from this research. I would prefer to think of them as dimensions within which to place a system, as anyone thinking of applying these should focus first on their particular problem and adapt as necessary. I find it best to summarise the lessons' in three key points to consider when designing a system like the VP:

**Presentation:** Does it look right, does it fit the accepted norms of the conferencing system (e.g. fonts, styles, etc), do the users understand what it really is (e.g. AA vs UB)?

**Interaction:** Will the users interact with it (e.g. public vs private), and can they interact with it in a straightforward and intuitive way (e.g. it is part of the system and doesn't require complex cutting and pasting)?

**Sustainability:** Can you maintain the content (e.g. do users give feedback and can you capture more content from the discussions)?

The next section looks at some practical requirements for Organisational Memories, developed recently by Abecker *et al.* (1998) and how they map onto this work. It is also clear that the major problems have not been of the technical kind but rather of the acceptance by the users, and I will come back to this issue later in section 8.3.

## 8.2 Practical requirements for organisational memories

Independently of this research Abecker *et al.* (1998) have developed a number of practical requirements for organisational memories. They also champion the need for "computer-assisted knowledge capitalisation" using assistant systems, similar in some ways to the Virtual Participant. They argue, with the use of Fredrick Brooks' (1996) co-operation formula: IA > AI, that the combination of an assistant system with the user (intelligence-amplification) is much more powerful than any separate AI system. The Virtual Participant fits with this argument.

Abecker *et al.* use their experience to specify a set of five requirements for the success of organisational memories:

1. Collection and systematic organisation of information from various sources
2. Minimisation of up-front knowledge engineering
3. Exploiting user feedback for maintenance and evolution
4. Integration into existing work environment
5. Active presentation of relevant information

These requirements have grown out of their experience in successful industrial practice. These guidelines (before they were known) have all been followed in the development of the Virtual Participant. Taking each point in turn it is possible to see how the Virtual Participant has addressed these requirements. The following sections then go on to look at some of the specific aspects of the Virtual Participant's design, at what has been learned, and at how other systems have handled similar problems.

1. **Collection and systematic organisation of information from various sources.** The course that the Virtual Participant is used on is typical of many distance education courses and involves a variety of media and materials. The students are provided with course texts, books, videos, audio cassettes, electronic conferencing, and tutorials. This makes up a complex mixture of multi-media and multi-modal resources for the students, and for the Virtual Participant. Although the initial prototypes have only made use of the text-based resources (which are already pretty wide-ranging and varied) an advanced system would wish to make use of other resources.

2. **Minimisation of up-front knowledge engineering.** In the development of the Virtual Participant there was very little access to an expert, requiring the minimisation of up-front knowledge-intensive engineering work. In a way this is a good thing; for this kind of system it is a significant advantage if it can be rapidly applied to other courses.

3. **Exploiting user feedback for maintenance and evolution.** The first prototype did not make any allowances for user feedback, although this was rectified in the second version. This feedback would enable the system not only to identify whether it has matched information to the correct discussion, but also whether that information is actually of any help. Unfortunately in this study the feedback turned out to be of little use, and few students provided it. In addition to this, in principle the system should be able to monitor the conferences to identify discussions about subjects of which it has no knowledge, and hence store that knowledge for future re-use.

4. **Integration into existing work environment.** The Virtual Participant behaves like any other user of the conferencing system. This removes the need for the users to run any specialised software, and minimises the amount of learning needed to use the system. With the second prototype the VP was also able to use formatting in the same way as normal users.

5. **Active presentation of relevant information.** Relevant information needs to be presented at a time and in a context where it will be useful. This is one of the shortcomings of FAQs where the user first has to recognise that the FAQ might contain the answer they were looking for, then they have to find the FAQ, and only then can they begin to look for the answer. The Virtual Participant contextualises its information to the current discussion, helping the students gain information that perhaps they did not even know they needed.

These five points cover key aspects which have helped the success of the Virtual Participant. The fact that these issues have also been identified independently as practical requirements for systems designed to support the sharing and reuse of knowledge confirms the value of these findings. Particularly relevant are requirements 2, 3, and 4 which have proven key not only to the development of the VP but can also be seen in Ackerman's work on another knowledge sharing system, Answer Garden (Ackerman, 1994a).

The next section looks at some of the reasons for the non-acceptance of knowledge management systems and looks at how the problems of usability and acceptance of the VP map onto them.

## 8.3 Problems with usability and acceptance

Knowledge management is a term which has recently become widely used, and often abused. The Virtual Participant itself is not only a system for knowledge sharing and reuse, but exists within a groupware system used to support collaborative learning, FirstClass. Knowledge management, sharing, and reuse are themselves group activities. From this position it seems likely that previous work by Grudin (1994) on the main problems for developers of groupware systems may be applicable. In his paper Grudin sets out eight problems facing developers:

1. **Disparity in work and benefit.** People may have to do extra work for little or no benefit; when this happens, the system will often fail. This is the problem of "role conflict" (Watt, 1993).
2. **Critical mass problems.** Some group applications really only work when a 'critical mass' of people use them.
3. **Disruption of social processes.** Group applications may break existing social rules and roles within an organisation or institution.
4. **Exception handling.** Group interaction is very complicated, and a lot of repair and improvisation may happen; applications need to be able to accommodate this.
5. **Unobtrusive accessibility.** In group interaction, some infrequently used features are still very important (e.g. privacy settings) and must always remain accessible.
6. **Difficulty of evaluation.** Because of the number of people involved, and the cultural embedding of the interaction, it is very hard to evaluate collaborative systems properly.

7. **Failure of intuitions.** Intuition is not a reliable guide when conceiving or designing any system, although it works rather more often for single-user systems.
8. **The adoption process.** Group systems need to be introduced into workplaces much more carefully than single-user systems.

With the Virtual Participant each of these problems has surfaced to some extent or another. Problem 6 has been particularly burdensome with the difficulty of assessing the impact of course changes and conferencing changes separately from the (quite large number of) software changes. The main problem yet to overcome is problem 1, in that it has proved very difficult to encourage contributions back into the system, although this problem may be starting to give with the significant changes in the tutor's acceptance of the system seen in the interviews discussed in chapter 7. Simply using conferencing for teaching (especially with the introduction of the VP) has suffered from problems 3, 7 and 8. Conferencing has been introduced into the OU's traditional distance teaching approach, but to date it has not been adopted by all courses. This in itself is a disruption of the existing social processes of teaching and learning. The introduction of the VP has had similar effects on the social processes being established in conferencing. Finally a number of the initial decisions made in the first version of the Virtual Participant were based on my own (and other people's) intuitions which have been found not to be a completely reliable guide - a good example was the choice of name.

Now let's look back to chapter 2 and the discussion of Chan's (1995) work on the sort of roles that pedagogical agents might take: artificial teacher, artificial learning companion, students' assistant, and tutors' assistant. The view reflected in these categorisations, and in many teaching systems, is of a one-to-one interaction between student and system. This reflects an obvious simplifying assumption which is mostly there so that a system can attempt to track what a student has learnt.

This section's emphasis on groupware is because I believe that we should be moving to the 'next level', so to speak. I argue that although single user systems have been shown to help teaching and learning we must also begin to look further at ways of supporting group interactions, like the Virtual Participant. Because of this it is necessary to learn from Grudin's past lessons above, and from Abecker *et al.'s* practical requirements as well as from the lessons learned from designing the VP. However, perhaps most importantly, to enable this work to be taken forward it is necessary to get over the key problem of acceptance.

The following are some representative issues in the non-acceptance of a knowledge management system and fall into three distinct categories (Masterton & Watt, 1999). The rest of this section presents those categories showing which traps the VP fell into, and which were intentionally avoided.

- **Task related problems.** When the system makes tasks seem harder than the user believes they should be, perhaps adding to a person's daily load, and provide a value worth less than the perceived effort required, then people will often stop using the system if they feel they can.
- **Culture related problems.** When the use of the system doesn't fit with the culture of the organisation, or of the individuals using the system, this may put people off using the system.
- **Individual related problems.** When personal space is being infringed (or is perceived as being infringed) in one way or another, people will often stop using the system.

For the Virtual Participant all these kinds of problem occurred, and were addressed in the following way:

*Task related problems*

1. **'I need special software.'** One original idea for the VP was to produce a personalised version of an assistant program which students could ask for help. Producing reliable software which I could distribute to all students and which would work on all platforms would have been a very difficult problem in itself, especially given the time available. However the main reason for ruling it out was because of the problems involved in convincing students that special software like this might help them, and then getting them to learn how to use it. This is also a culture related problem.
2. **'I have enough to do each day as it is.'** This was not a problem I attempted to address directly; what I was trying to do was to integrate the VP into the students' everyday working pattern. The way I achieved this was to have it available through the conferencing system behaving like 'any other user', the idea being that they could chose to ignore messages from it if they so wished. This design decision also addressed this problem.
3. **'I can't find anything useful anyway.'** As stated earlier, the real benefit of the Virtual Participant comes from contextualising the contents of the archive to the

current discussion. By using the available information in this way, I hoped that the VP's knowledge base would be as useful as possible, without losing relevance.

4. **'I don't know how to do the task very well.'** No matter how simple you try to make things there will always be those who fail to understand - or even read - the instructions. By making the task of getting further information from the VP similar to the task of normally replying to messages we hoped to eliminate this problem. However, even this did not cover all eventualities as the VP received a number of blank e-mails from students using old versions of the software.

*Culture related problems*

1. **'I need to change the way I work.'** This is a problem we experienced both with tutors and students, although for different reasons. At first, the tutors felt at best the VP was just one more thing to worry about, and at worst it would interfere with the teaching process. For the students the VP was something they had never come across before, requiring a change from their normal conferencing approaches. This was especially true in the first study where those who preferred to remain lurkers were required to post public messages to get more information. As we concluded after the interviews in chapter 7, the VP could only be successful if the tutors incorporated it into their teaching approach. It is likely that this would also lead to greater student acceptance. This is a typical culture related problem, which is also mitigated by ensuring, as the VP does, that no new software need be used.
2. **'I don't think this system is right for this organisation.'** If we insert the word 'teaching' before the word 'organisation' we can summarise some of the objections we initially received to the VP. Some tutors felt that the approach championed by the VP was against the culture of their current teaching methods. This can only be overcome by convincing them of the benefit of using the system. As this study shows, to a large extent this is indeed possible.

*Individual related problems*

1. **'I feel my privacy is being invaded.'** This problem was the specific reason why the contents of the VP knowledge base were anonymised. During the study the students were not, generally, asked if they minded having their messages incorporated into the knowledge base. Because of the nature of the study it was rarely the same students in each year and so it was only the tutors who had their

words 'come back to haunt them', so to speak. However privacy is a complex issue, and one which cannot be fully dealt with in this thesis. However, in those cases where I have asked students if they mind having their words incorporated into the system they have often expressed surprise and are rather flattered.

2. **'I'm shy.'** This is an issue that is common to the world in general and cannot be ignored. In the first prototype students had to ask questions publicly. Although in the questionnaire they responded in the majority that this was fine their behaviour indicated otherwise. Adding the ability to ask questions privately, and making this the main method of interaction for the second prototype led to dramatically increased use of the VP. Unfortunately, as discussed in chapter 7, those who asked questions were still only the same users who participated in the discussions.
3. **'I feel the system is not representing my views.'** This is not a problem specific to the VP as it is intended to present multiple possible viewpoints. The down side of this is that students may not find anything to help support their ideas, and tutors may find that statements within the system disagree with their preferred viewpoint. This is very important given that the attitude of some users was that because the material in the system was selected by the tutors then it must have been correct, see section 7.4.3.
4. **'I can't control the system.'** A major problem with the first prototype was the tutors felt that they had no control over the system. With the second prototype we had agreed to allow them complete control. However they decided that using this would require too much extra work and opted to deal with any control problems through me. This highlights the fact that perceived control was much more important than the actual control (Langer, 1983), sometimes known as 'the Langer effect'.

**Reflection on Social and Motivational Issues with the VP**

Looking back to section 3.4 on the system rationale for the first prototype there were some politically-motivated factors behind the design. That section stated that "The aim of the exercise is to win over the hearts and minds of those whose environment it impinges on the most - primarily the electronic tutors, and secondarily the students". This statement was accepting, before the research even began, that the motivation of those involved is highly important. When Grudin (1994), talks about the introduction of new systems he advises the developer to "provide education that demonstrates a positive impact on the work". A clear example of this from the Virtual Participant study is in its use of feedback. There is no clear benefit to the student providing

feedback, but work from recommender systems (Resnick & Varian, 1997) indicated that a small number would happily do so. With the second study when the students were asked about why they did not give feedback in practice, section 7.4.3, many said that they felt that it was too much effort or of little benefit. For those that did give feedback as there was no response to their efforts they felt demotivated. An important lesson from this is that if natural motivation exists then it is important to support and channel it.

When describing the categories of common reasons for non-acceptance of knowledge management systems the result was an overlapping set of task, culture, and individual related problems. Problems related to the task are associated with usability and integration, and can mostly be addressed by improving usability and integration. However, culture and individual related problems are mostly of a social nature. For many students, using electronic conferencing in their learning activities is a new experience. It is a group process that takes place asynchronously and at a distance, completely the opposite of the synchronous, co-located, experience of tutorials. Some argue that for students to be successful in such an environment it is necessary to socially acclimatise them to that environment (Salmon & Giles, 1996). Once acclimatised they are as aware of the social norms in the conferencing environment as those in the face-to-face environment. In this situation it is the tutors who set what will be considered the social norms. By convincing the tutors of the value of systems such as the Virtual Participant you address any fears of "I don't think this system is right for this organisation". It seems probable that once the tutors are convinced, and have adopted the system into their conferencing norms, the students will follow.

**Summary**

This section has provided a more detailed view on the usability and acceptance problems of systems like the VP. Starting with the work of Grudin on the problems facing the designers of groupware systems, we then moved on to look at some of the non-acceptance problems facing knowledge management systems. At all times the relationship between these problems and the experiences with the Virtual Participant have been made explicit. Between discussing Grudin's work and the non-acceptance problems, I highlighted what I see as the need to move forward from a traditional view of agents in learning systems towards the approach pioneered by the Virtual Participant. The next section takes this argument forward, first by looking at the more recent roles for agents, specifically in knowledge management systems, and then at ways of designing more effective pedagogical agents.

## 8.4 Agents in learning systems

This section comes in two parts. The first looks specifically at the roles an agent might adopt in a learning system, drawing on examples. The second then focuses on the design of pedagogical agents, tying the issues of usability and acceptability together with the role of the system.

### The roles of agents in learning systems

The previous sections argue strongly for the need to consider the wider scope of social issues in agent-based learning systems. In this section we look at some of the common roles adopted by knowledge management systems (Masterton & Watt, 1999). It is necessary to be aware of the diversity of possible roles when designing systems, and I feel that these are equally applicable to approaches such as the Virtual Participant. By considering in detail how the system will present and fulfil its role, this can help to plan the approach better .

Some of the common roles adopted by knowledge management tools and systems are:

- Knowledge management systems as assistants e.g. (Abecker et al., 1998)
- Knowledge management systems as matchmakers e.g. (Foner & Crabtree, 1996)
- Knowledge management systems as librarians e.g. (Watt, 1998)
- Knowledge management systems as reporters e.g. (Domingue & Scott, 1998)
- Knowledge management systems as editors e.g. (Domingue & Scott, 1998)
- Knowledge management systems as oracles e.g. (Ackerman, 1994)
- Knowledge management systems as critics e.g. (Fischer *et al.*, 1993)
- Knowledge management systems as bards e.g. (Masterton, 1997; Masterton, 1998)
- Knowledge management systems as village gossips (Krulwich & Burkey, 1996; Krulwich & Burkey, 1997)

Answer Garden (Ackerman, 1994) and similar systems have been described as 'oracles', because, like the Delphic Oracle, people could go to them to get relatively definitive answers to questions they might have. We have described the Virtual Participant as a 'bard', because, to put it simply, it re-told stories from the past. We

have described InfoFinder and ContactFinder (Krulwich & Burkey, 1996; Krulwich & Burkey, 1997) as like 'village gossips' because they passed on short snippets of information that they believed you might be interested in, without this information being solicited. In teaching situations students and teachers can also use assistants, librarians, and critics to help guide them in learning and information discovery. The roles presented here may not all be appropriate for use in an educational setting, but that really depends on the approach that you are looking to take. It is clear that no single role, even that of the "teacher", is adequate to support the whole learning process.

There are a wide variety of different social roles in the learning process. Because of this diversity the cultural associations of each role will be very different and this makes them very difficult to compare. The roles which a system adopts, the problems with usability and acceptance, and the dimensions which we identified as particularly important for the Virtual Participant all come into play in the design process.

**Designing effective pedagogical agents**

Based on the evidence in this thesis, it seems likely that successful collaborative learning systems will have some features of agent-based systems. The main example has been the Virtual Participant, from which we have already learned a number of important lessons. By combining these lessons with the evidence from other systems and the idea of the different social roles adopted by different systems, we can look at developing a few informal guidelines for the design and introduction of pedagogical agent systems.

There are two key aspects to consider in the introduction of agent systems. First is usability, can people use the system? Second is acceptability, do people use the system? As far as usability is concerned there is much basic research in the field of human-computer interaction which can help achieve this (e.g., Norman, 1988). Acceptability is frequently a more problematic issue for groupware and collaborative learning systems alike, (Grudin 1994) contains a number of useful strategies and guidelines which may be of help.

The problem we have here is the introduction of agent-based systems. Papers like (Norman, 1994) and (Erickson, 1997) look at some of the social issues raised by this kind of system. As we have argued earlier, these problems are of particular importance for a system like the VP, and the roles which the system is to adopt need to be carefully considered. Matching the agent teaching system to existing accepted

roles shows some promise as a possible technique to facilitate its acceptance, and ensure that it is properly designed to be acceptable.

Earlier I looked at the important design dimensions which had arisen. These dimensions form part of what we can call the system's 'character space'. As we found with Uncle Bulgaria and the Active Archive, placing the system in the right place on the anthropomorphism scale can greatly effect user opinion. However, over and above the four dimensions identified in this research, anthropomorphism, privacy, visibility, and flexibility, it is likely that there are others yet to be identified.

Placing your system in the right place in this dimensional space is strongly dependent on the roles you wish your system to adopt. For example, you might want an integrated system which plays a number of different roles and therefore appears differently to the users when playing each of those roles. By balancing the trade-off between different dimensions (e.g. how private the system is *versus* how visible all the information is) depending on the chosen role, you have the tools you need to win at least half the battle.

When the design is complete the other half of the battle is in the marketing. When introducing a system you need to think through how you are going to 'sell' it to the users, even if they are students and tutors. The way a user views a system in flyers and manuals (and even teaching materials) can have a significant effect on how they perceive the system. Specifically shying away from calling any system an agent-based system is likely to be a good move. Grudin (1994) also provides guidelines for "managing acceptance", which start by clearly identifying the group's problems when designing the system, carefully introducing the system, explaining its benefits, and providing "education that demonstrates a positive impact [of the system] on the work". All these techniques can help to make systems like the VP easier to introduce successfully.

## 8.5 Revisiting the goals of the Virtual Participant

The purpose of this section is to revisit the issues raised by the Virtual Participant studies and reflect on those goals I set back in section 1.3. To that end, the following four sub-sections review each goal, asking a number of questions, and looking at the evidence available to answer them. After reviewing the goals I make a special point of revisiting the idea of stories, the power of which have not fully been explored in this work.

- **The required effort of tutors can be reduced.** The VP will aim to provide content input relevant to the context of discussion - reducing the need for tutor involvement.
- **Discussion can be stimulated.** Discussion is known to be an important element of learning and because of this the VP will aim to stimulate more discussion by its input into existing discussions and through posing questions. This will improve the value of the electronic conferencing and is linked to the next point.
- **Similar discussions take place each year.** With similar course materials and assignments each year it is likely that similar discussions will take place. This study will not only show that this is correct, but also that previous years' discussions can prove an important learning resource improving the experience for current students - especially when contextualised to their discussion.
- **It is possible to help students even if they don't participate.** Not all students like to take part in discussions, often up to two thirds will lurk. The VP will aim to provide more quality content for lurkers and raise the quality level of discussions to encourage lurkers to participate. The VP will also interact with them directly, encouraging them to contribute.

## The required effort of tutors can be reduced.

**Did the VP reduce the need for tutor involvement?** The VP set out to provide content input relevant to the context of the discussion which was taking place. The aim of this was that the extra material would prove useful to the students in solving their problems, which in turn would reduce the need for tutor involvement. Based on the questionnaire feedback, feedback to VP messages, and reading the discussions, the VP did succeed in adding relevant input. Having no quantitative measures of tutor involvement it is impossible to say overall whether there were any significant differences from what would have been normal for the studies. However by analysis of those threads which the VP was involved in, especially in the second study, it can be seen that the VP's input probably required the tutors to provide extra content. This extra content was provided in both instances where a student replied to a VP message asking for more information. There is also the example of the additional thread of discussion started by the tutors in response to some of the VP's messages, discussed in section 7.2.2.

**Did the tutors accept the system?** Based on the data it is clear that the VP did not reduce tutor involvement, and in fact actually increased it. I do believe that in these instances the additional information from the tutors added value to the students' learning experience. I also think that it is important to remember that the tutors had come to accept the system, relaxing their wish for control, and they seemed quite happy with it in their feedback. Being able to build on this buy-in could increase the effectiveness of the system as an assistant for the tutors, perhaps through role extensions like welcoming people when they first login and raising discussion items.

### Discussion can be stimulated.

**Did the VP actually stimulate additional discussion?** As noted in answer to the previous points there are clear examples in the second study where additional discussion was stimulated because of the messages from the VP. This indicates that a system like the VP can help to achieve this goal. Although, in this study, there was an increased burden on the tutors because of the additional discussion it is important to remember that the VP operated separately to the tutors. If the content and interactions of a VP system could be linked with the teaching approaches, specifically to assist the tutors, then it seems probable that systems like this can effectively promote discussion and through this enhance students' learning.

### Similar discussions take place each year.

**Did students actually learn?** During the studies threads in different years were given the same title (see section 4.5) and one student posted their first contribution in response to the VP posting a message from a previous year about an identical problem to the one the student was experiencing (see section 7.2.2). These observations clearly demonstrated that there were repetitions in the discussions between years. What can not be concluded, at least quantitatively, is whether the students learnt from the VP's contributions. The evidence from questionnaire feedback, students asking for more details about VP contributions publicly and privately, and the large amount of reading of contributions and the archive all point to learning having taken place and therefore value having been gained from these contributions.

### It is possible to help students even if they don't participate.

**Were lurkers encouraged to participate?** The second study provided an excellent example where a student, who had already asked private questions of the VP, contributed to the discussion in direct response to one of its messages (see section

7.2.2). This is important because it was the first public contribution to the discussion by this student. The fact that they had already asked questions is important because it indicates that they had no illusions about the nature of the Virtual Participant. The message they chose to respond to was about a student in a previous year who worked in a similar firm and had a very similar problem. This demonstrates how close the context of the VP messages can be to the problems experienced by students, so even if they are not totally relevant to the discussion they can still prove useful. It is important to remember that this message from the VP was also publicly available in the archive. There is a second example of stimulated discussion (also in section 7.2.2) where the student could not be classified as a lurker as they had already contributed to conference. However, this first contribution had been just 3 hours before and this was during the second TMA, so they had had ample opportunity to contribute previously.

**Did students learn vicariously?** Using this same evidence on learning from section 8.5.3 above, especially that of students reading the archives, there are important indications that vicarious learning has taken place. This is inferred from the observation that there were a large number of students who read many of the AA messages in the archive, but never any in the discussion conference - it seems probable that they would not have committed that time and effort if they were not getting value from the messages. Although there were no clear cut cases of students who did not contribute to the discussions asking the VP questions there is an example (section 7.2.1) of a student whose only contribution was to ask a question publicly, which was probably an accident. If this question had been asked privately it would clearly be an example of vicarious learning.

## The Power of Story

The Open University Business School courses teaches students to become "Reflective Practitioners" (Schön, 1983) as a process of sensemaking (Weick, 1995) to understand the practice of what an MBA teaches them. If we think about students who come together to learn about a course as a form of organisation then they use stories to help build shared language, mindsets, and understanding. Weick (1995, p129) presents a non-exclusive list of seven functions of story telling:

- Stories aid comprehension because they integrate that which is known about an event with that which is conjectural
- Stories suggest a causal order for events that originally are perceived as unrelated and akin to a list

- Stories enable people to talk about absent things and to connect them with present things in the interest of meaning
- Stories are mnemonics that enable people to reconstruct earlier complex events
- Stories can guide action before routines are formulated and can enrich routines after those routines are formulated
- Stories enable people to build a database of experience from which they can infer how things work
- Stories transmit and reinforce third-order controls by conveying shared values and meaning

Weick also sees stories as having features such as plausibility, coherence, memorability and resonance, and so forming key vehicles for retaining plausible truths about the world. Linking this with his process of sensemaking then stories become the method of remembering. Weick's process consists of the following three steps:

**Enactment:** The process of actively creating the environment

**Selection:** The process of generating answers to questions

**Retention:** The process of storing the output of successful sensemaking

This seems to bear some similarity to Kolb's learning cycle, figure 1.2, where Enactment maps onto Active Experimentation, Selection onto Concrete Experience, and Retention onto Reflective Observation. However it seems that Abstract Conceptualisation overlaps with both the Enactment phase, as a prologue to sensemaking, and the Retention phase, as the epilogue to sensemaking.

By taking "...the idea that people think narratively rather than argumentatively or paradigmatically" (Weick 1995, p127) combined with the seven functions described above there are implications as to the power stories could have in teaching. The value stories provide is to link groups of students across years, all of whom are learning about the same theories, functions, frameworks, case studies, etc. They can also help the tutors engender shared mindsets within the students through the management of their learning experience as they attempt to relate what they are learning to their practical everyday problems. If we consider the learning organisation of the course across time then the stories collected provide an organisational memory of the course. Storing these stories and making them available for future organisational members

(the students) can only serve to improve the learning experience by facilitating the sensemaking process.

> In organisations, storytelling is the preferred sense-making currency of human relationships among internal and external stakeholders. People engage in a dynamic process of incremental refinement of their stories of new events as well as on-going reinterpretations of culturally sacred story lines. When a decision is at hand, the old stories are recounted and compared to unfolding story lines to keep the organisation from repeating historically bad choices and to invite the repetition of past successes.

(Boje, 1991, p. 106)

It seems clear from this is that there is still a lot of potential from the approach of capturing and telling stories in a teaching environment, and that perhaps recent advances in the domain of organisational science can be put to good use in the design of systems like the VP. It is also clear that the full value of story-telling in the context of learning as a sensemaking process was not fully realised in these studies, and should be analysed more fully in the future.

## 8.6 Vision of the future

In chapter 1 I originally proposed combining various existing technologies to create the Virtual Teaching Assistant (VTA). As the name implies, the role proposed for this kind of system was that of a teaching assistant. Continuing with the approach championed with the Virtual Participant, it would not be intended to replace tutors, but rather to extend the support available to them for the benefit of their students. McKillop (1997) has found that the more guidance and structure students are given in support, rather than being required to follow, proves beneficial to their studies. This confirms the experience of large distance teaching organisations, like the Open University, that have found that the support provided by a time table and key way points that the students should achieve during their studies are of critical importance in maintaining high retention, completion, and satisfaction rates amongst their students (Daniel, 1996).

The goal of the VTA would be to integrate such technologies to the benefit of both tutors and students. This could be a multi-agent system in which course guides provided for the students (Sumner & Taylor, 1997) are linked to the work we have done with the VP. To date each of these technologies has been implemented

separately, but by combining them it would be possible to move towards a more uniform presentation of materials. By additionally integrating the VTA and multi-media course materials, a whole new range of possibilities opens up. Steps in this direction are already being taken in projects like the Internet Science Visualisation Laboratory (Mulholland & Domingue, 1998). Here the student is able to submit to their tutor example traces of Prolog programs they are having trouble with. The tutor can view these, add text and audio annotations and return them to the student. The VTA could fulfil an interface role for systems like this, perhaps even storing case-bases of previous interactions which might prove useful to the students.

The Virtual Teaching Assistant is only one vision of a future of many possibilities. I strongly believe that the development of systems to support groups of learners is of fundamental importance. I also believe that more integrated approaches are necessary to advance the use of technologies such as these.

But what would be the next stage of development for the current Virtual Participant? One obvious area for development are some of the role extensions mentioned already. In the current version none of the controlling or notification mechanisms for tutors were implemented, only a way of delaying the sending of messages. These features would be very important if tutors are going to be more proactive in their use of the system, especially in more prescriptive courses where the tutors or course team might want certain information to be made available only at certain times. This would be necessary so that those involved in delivering the course can be satisfied that there would not be any conflicts with their design. By using the VP more as an assistant they may also choose to use it to greet people and introduce topics of discussion at appropriate times. Another restriction of the current implementation of the VP is that there is no authoring environment for adding, deleting, and changing the contents of the existing knowledge base. In the long run it must be possible for naive users to be able to do this. This is especially true if the students are to be allowed to add information themselves, even if mediated by the tutors.

One technological shortcoming of the current implementation of the VP has been in the matching between current discussions and the information stored in the knowledge base. The way forward here is to work on a more sophisticated natural language system. One system which might prove effective is 'Kenmore' (Cardie, 1994), a case-based natural language understanding system which could be adapted to index the past discussions and identify matches with current ones. It could be trained on the existing course text which would also provide a way of pointing students to relevant sections. A related option is to introduce some of the ideas from spreading

activation networks where the case-base would be structured in such a way that the matching of some stories made related ones more likely to be matched and suppressed the more unlikely ones.

With the students wishing to have an easier structure to browse and a better way of asking questions of the system, perhaps this would be an ideal application area for Answer Garden (Ackerman, 1994). One of the problems of Answer Garden was allocating experts to knowledge areas, but in our case we already have experts (i.e. tutors) allocated to students. In this way when a students cannot find an answer they could be directed to contact their tutor; the tutor could add new information to the knowledge base, and could also share important points within their tutor group, or with the whole course in general. Coupled with the VP approach answers from a system similar to Answer Garden could be made directly available when useful. If students were interested they could then use it to search the surrounding questions, perhaps through an additional mechanism for asking more questions and maintaining the knowledge base.

Finally the work already started in section 7.3 could lead to some more quantitative measures of how discussions progress. If these measures could be linked to the matching and delaying systems then it may be possible to improve the timing of the contributions of the Virtual Participant for when the discussion has lagged, or questions have been ignored.

Taking these suggestions forward as ways to improve the technology of the system, but perhaps more importantly, building on the will of the tutors to contribute seems likely to enable the Virtual Participant to achieve even greater success.

## 8.7 Epilogue

This thesis has shown the possibilities of integrating theories of learning cycles and supported learning conversations successfully into a tertiary courseware system. Through this work guidelines have been developed for usability, introduction, the types of roles such a system may take, and the use of agents in teaching systems. This work has only just begun to explore the power of story telling in computer supported cooperative learning and this is certainly something I would like to see explored further.

The VP project has clearly demonstrated that:

- Similar discussions do indeed take place in each year and so past discussions can be reused to help current students.
- It is possible to stimulate more discussion by introducing information relevant to the current discussion.
- It is possible to draw lurkers into discussions, and that they will also learn vicariously from the additional information available.

However the project did not manage to demonstrate a reduction in the effort of the tutors, and there is evidence that perhaps it was increased slightly. In future incarnations systems such as this need to be fully integrated into the teaching programme to truly demonstrate their potential.

# Bibliography

Abecker, A., Bernardi, A., Hinkelmann, K., Kühn, O., & Sintek, M. (1998). Toward a Technology for Organizational Memories. *IEEE Intelligent Systems and their Applications*, **13**(3), pp. 40-48.

Ackerman, M. S. & Malone, T. W. (1990). Answer Garden: A Tool for Growing Organizational Memory. In *Proceedings of ACM Conference on Office Information Systems*, (pp. 31-39). Cambridge, MA, USA

Ackerman, M. S. (1994a) *Answer Garden: A tool for Growing Organisational Memory*. PhD Thesis, Massachusetts Institute of Technology.

Ackerman, M. S. (1994b). Augmenting the Organisational Memory: A Field Study of Answer Garden. In *Proceedings of ACM Conference on Computer Supported Cooperative Work.*

Ackerman, M. S. & McDonald, D. W. (1996). Answer Garden 2: Merging Organisational Memory with Collaborative Help. In *Proceedings of ACM Conference on Computer-Supported Cooperative Work.*, pp97-105

Beresford, E. (1968). *The Wombles*. London: E. Benn.

Boje, D. M. (1991). The storytelling organisation: a study of story performance in an office-supply firm. *Administrative Science Quarterly*, **36**, 106-26.

Boyce, A. J. (1969). Mapping Diversity. In *Proceedings of Colloquium in Numerical Taxonomy*, University of St. Andrews.

Brooks, F. P. (1996). The Computer Scientist as Toolsmith II. *Communications of the ACM*, **39**(3), pp61-68.

Cardie, C. (1994) *Kenmore: A Case-Base Nominal Disambiguation System*. PhD Thesis, Yale.

Chan, T.-W. (1995a). Artificial Agents in Distance Learning. *International Journal of Educational Telecommunications*, **1**(2/3), 263-282.

Chan, T.-W. (1995b). Social Learning Systems: An Overview. In B. Collis & G. Davies (Eds.), *Innovative Adult Learning with Innovative Technologies* (pp. 31-47). Elsevier Science B.V. (North-Holland).

Conklin, J. & Begeman, M. (1988). gIBIS: A Hypertext Tool for Exploratory Policy Discussion. In *Proceedings of Computer Supported Collaborative Work (CSCW)*, (pp. 140-152). ACM.

Daniel, J. (1996). *The Mega-Universities and Knowledge Media.* London: Kogan Page.

Davies, N. J., Stewart, R. S., & Weeks, R. (1998). Knowledge Sharing Agents Over the World Wide Web. *British Telecom Technology Journal*, **16**(3), 104-109.

Domingue, J. & Scott, P. (1998). KMi Planet: A Web Based News Server. In *Proceedings of Proceedings of the Asia Pacific Computer Human Interaction Conference (APCHI'98)*, . Shonan Village Center, Hayama-machi, Kanagawa, Japan: 15-17 July.

Eisenstadt, M. (1993). Tales of Debugging from The Front Lines. In *Proceedings of Empirical Studies of Programmers V.*

Eisenstadt, M. (1997). My Hairiest Bug War Stories. *Communications of the ACM*, **40**(4), 30-37.

Eisenstadt, M., Buckingham-Shum, S., & Freeman, A. (1996). KMi Stadium: Web-based Audio/Visual Interaction As Reusable Organisational Expertise. In *Proceedings of Workshop on Knowledge Media for Improving Organisational Expertise, 1st International Conference on Practical Aspects of Knowledge Managment.* Basel, Switzerland

Eisenstadt, M., Price, B. A., & Domingue, J. (1992). Redressing ITS Fallacies Via Software Visualization. In E. Lemut, B. d. Boulay, & G. Dettori (Ed.), *Proceedings of Cognitive Models and Intelligent Environments for Learning Programming*, 111 (pp. 220-234). Genova, Italy: Springer-Verlag.

Erickson, T. (1997). Designing Agents as if People Mattered. In J. Bradshaw (Eds.), *Software Agents* (pp. 79-96). AAAI Press / MIT Press.

Ferguson, W., Bareiss, R., Birnbaum, L., & Osgood, R. (1992). ASK Systems: An Approach to the Realization of Story-Based Teachers. *The Journal of the Learning Science*, **2**(1), 95-134.

Fischer, G., Nakakoji, K., Ostwald, J., Stahl, G., & Sumner, T. (1993). Embedding Critics in Design Environments. *The Knowledge Engineering Review*, **8**(4), pp. 285-307.

Foner, L. (1993). What's An Agent, Anyway? A Sociological Case Study. In ftp://media.mit.edu/pub/Foner/Papers/Julia.Agents--Julia.ps, Agents Memo 93-01, MIT Media Lab.

Foner, L. (1996). A Security Architecture for Multi-Agent Matchmaking. In *Proceedings of The Second International Conference on Multi-Agent Systems*. Keihanna Plaza, Kansai Science City, Japan: December '96.

Foner, L. & Crabtree, I. B. (1996). Multi-agent Matchmaking. *British Telecom Technology Journal*, **14**(4), pp. 115-123.

Franklin, S. & Graesser, A. (1996). Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Proceedings of Intelligent Agents II - Agent Theories, Architectures, and Languages*, LNAI Vol. 1037. JCAI'95-ATAL Workshop, Montreal, Canada: Springer-Verlag.

Grudin, J. (1994). Groupware and Social Dynamics: Eight Challanges for Developers. *Communications of the ACM*, **37**(1), 92-105.

Hall, R., Miyake, N., & Enyedy, N. (1997). Proceedings of the Conference on Computer Supported Collaborative Learning. OISE, University of Toronto, Toronto, Canada

Hambridge, S. (1995). *Netiquette Guidelines* (Internet Request for Comments No. RFC1855). Intel Corp.

Kay, A. (1990). User Interface: A Personal View. In B. Laurel (Eds.), *The Art of Human-Computer Interface Design* (pp. 191-207). Reading, Massachusetts: Addison-Wesley.

Koda, T. (1996) *Agents with Faces: A Study on the Effects of Personification of Software Agents*. Masters Thesis, MIT.

Kolb, D. A. (1984). *Experiential Learning: Experience as The Source of Learning and Development*. New Jersey: Prentice Hall.

Kolodner, J. (1993). *Case-based Reasoning*. Morgan Kaufmann, Los Angeles, ISBN: 1558602372.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*, **40**(3), 77-87.

Krulwich, B. & Burkey, C. (1996). The ContactFinder agent: Answering bulleting board questions with referrals. In *Proceedings of AAAI 96, The Thirteenth National Conference on Artificial Intelligence, The Eighth Conference on Innovative Applications of Artificial Intelligence*, Volume One (pp. 10-15). Portland, Oregon, USA: AAAI Press.

Krulwich, B. & Burkey, C. (1997, September/October). The InfoFinder Agent: Learning User Interests through Heuristic Phrase Extraction. *IEEE Expert*, p. 22-27.

Langer, E. J. (1983). "*the PSYCHOLOGY of CONTROL.*" Beverly Hills: SAGE.

Lanier, J. (1995). Agents of Alienation. *Interactions*, **2**(3), 66-72.

Lanier, J. & Maes, P. (1996, 15-24 July). Intelligent Agents = Stupid Humans. *HotWired*, http://www.hotwired.com/braintennis/96/29/index0a.html.

Lee, J., Dineen, F., & McKendree, J. (1997). Supporting Student Discussions: It Isn't Just Talk. In D. Dicheva & I. Stanchev (Ed.), *Proceedings of IFIP WG3.3 Research on Educational Applications of Information Technologies*, (pp. 124-136). Sozopol, Bulgaria: Virtech Ltd, Sophia.

Luger, G. & Stubblefield, B. (1997). *Artificial Intelligence*. Addison-Wesley Longman.

Maes, P. (1994). Agents that Reduce Work and Information Overload. *Communications of the ACM*, **37**(7), 31-40.

Martin, J. & Henry, J. (1993). *B882 - Creative Managment.* The Open University.

Masterton, S. (1995) *Modeling Changing Interests in the Domain of Electronic News.* MSc. Thesis, The University of Edinburgh.

Masterton, S. (1997). The Virtual Participant: Lessons to be Learned from a Case-Based Tutor's Assistant. In *Proceedings of Conference on Computer Supported Collaborative Learning,* Toronto, Canada:

Masterton, S. (1998). Computer Support for Learners Using Intelligent Educational Agents: The Way Forward. In *Proceedings of International Conference on Computers in Education,* Volume 1, 211-219, Beijing, China.

Masterton, S. & Watt, S. (1999). *Oracles, Bards, and Village Gossips, or, Social Roles and Meta Knowledge Management* (Techreport No. KMi-TR-82). The Open University, Knowledge Media Institute.

Mayes, T. (1995). Learning Technology and Groundhog Day. In W. Strang, V. Simpson, & D. Slater (Eds.), *Hypermedia at Work: Practice and Theory in Higher Education* Canterbury: University of Kent Press.

Mayes, J. T. & Neilson, I. (1995). Learning From Other People's Dialogues: Questions About Computer-Based Answers. In B. Collis & G. Davies (Eds.), *Innovative Adult Learning with Innovative Technologies* (pp. 31-47). Elsevier Science B.V. (North-Holland).

McCarthy, J. (1979). *Ascribing Mental Qualities to Machines* (Technical Report Memo No. 326). Stanford CA: Stanford AI Lab.

McKendree, J. (1996). The Vicarious Learner: Rethinking the Use of Dialogue in Teaching and Learning. In *Proceedings of Issues in Interactive Multimedia,* Workshop, New Mexico: http://www.hcrc.ed.ac.uk/gal/vicar/VicarPapers/VicarLearn.html

McKendree, J. & Mayes, J. T. (1997). "The Vicarious Learner": Investigating the Benefits of Observing Peer Dialogues. In *Proceedings of Computer-Assisted Learning,* pp161-164, Exeter, UK

McKendree, J., Stenning, K., Mayes, T., Lee, J., & Cox, R. (1997). Why Observing A Dialogue May Benefit Learning: The Vicarious Learner. In *Proceedings of the 8th PEG Conference*, pp156-163, Sozopol, Bulgaria

McKillop, C. (1997) *On-line Study Guides for Distance Education Students: Can 'Advisor' Agents Help?* MSc Thesis, University of Sussex.

Metral, M. E. (1993) *Design of a Generic Learning Interface Agent.* MSc Thesis, Massachusetts Institute of Technology.

Miller, G. A., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behaviour.* New York: Holt.

Mulholland, P. & Domingue, J. B. (1998). Sharing Programming Knowledge Over the Web: The Internet Software Visualization Laboratory. In M. Eisenstadt & T. Vincent (Eds.), *The Knowledge Web* , pp185-197, Kogan Page.

Newell, A. (1982). The Knowledge Level. *Artificial Intelligence*, **18**(1), pp87-127.

Nishimoto, K., Sumi, Y., & Mase, K. (1996). Towards an Outsider Agent for Supporting a Brainstorming Session - An Information Retrieval Method from a Different Viewpoint. *Knowledge-Based Systems*, **9**, 377-384.

Nixon, T. & Salmon, G. (1995). Spinning Your Web: Interactive Computer-Mediated Conferencing - Its Potential for Learning and Teaching in Higher Education. In *Proceedings of The Annual Conference of the Society for Research Into Higher Education*, Edinburgh

Norman, D. A. (1988). *The Psychology of Everyday Things.* New York: Basic Books.

Norman, D. A. (1994). How Might People Interact with Agents. *Communications of the ACM*, **37**(7), 31-40.

Pask, G. (1975). *Conversation, Cognition and Learning: A Cybernetic Theory and Methodology.* Amsterdam: Elsevier.

Resnick, P. & Varian, H. R. (1997). Recommender Systems. *Communications of the ACM*, **40**(3), 56-58.

Salmon, G. & Giles, K. (1996). Training Virtual Management Teachers. In *Proceedings of Online Educa Conference*, Berlin, Germany

Salmon, G. & Giles, K. (1997). Moderating On-Line. In *Proceedings of Online Educa Conference*, Berlin, Germany

Salton, G. & McGill, M. (1983). *Introduction to Modern Information Retrieval.* New York: McGraw-Hill.

Schank, R. & Cleary, C. (1995). *Engines for Education.* Lawrence Erlbaum Associates, New Jersey, USA.

Schon, D. (1983). *The Reflective Practitioner: How Professionals Think in Action.* London: Basic Books.

Scott, P. J. & Phillips, M. (1998). Developing Web-Based Student Support Systems; Telling Student Stories on the Internet. In M. Eisenstadt & T. Vincent (Eds.), *The Knowledge Web* , pp79-96, Kogan Page.

Sheth, B. D. (1994) *A Learning Approach to Personalized Information Filtering.* MSc Thesis, Massachusetts Institute of Technology.

Smeaton, C. (1995). The AnswerWeb. In *Proceedings of Fourth International World Wide Web Conference "The Web Revolution",* Boston, Massachusetts, USA

Smeaton, C. & Slater, A. (1994). Integrating Simulations and WWW Courseware. In *Proceedings of the second International WWW conference,* Chicago, USA

Sokal, R. R. & Sneath, P. H. A. (1963). *Principles of Numerical Taxonomy.* San Francisco: W.H. Freeman.

Sumner, T. & Taylor, J. (1997). Coping with Virtuality: Steps towards a Personal Learning Manager. In T. O'Shea (Ed.), *Proceedings of UNESCO Conference on Virtual Learning Environments,* Milton Keynes, UK

Tansley, D. S. W. & Hayball, C. C. (1993). *Knowledge-Based Systems Analysis and Design.* Hemel Hempstead, UK: Prentice Hall.

Thomas, P., Carswell, L., Price, B., & Petre, M. (1998). A Holistic Approach to Supporting Distance Learning Using the Internet: Transformation, not Translation. *British Journal of Educational Technology*, **29**(2), 149-161.

Turkle, S. (1997). *Life on the Screen: Identity in the Age of the Internet.* Weidenfeld & Nicolson.

Watt, S. (1993). Role Conflict in Groupware. In *Proceedings of The First International Conference on Intelligent Cooperative Information Systems*, Rotterdam

Watt, S. (1996). Artificial Societies and Psychological Agents. *British Telecom Technology Journal*, **14**(4), 89-97.

Watt, S. N. K. (1998). Psychological Agents and the New Web Media. In M. Eisenstadt & A. T. Vincent (Eds.), *The Knowledge Web* , pp239-248, Kogan Page.

Weick, K. E. (1995). *Sensemaking in Organisations*. London: Sage.

Whalley, P. (1995). Imagining with multimedia. *British Journal of Educational Technology*, **26**(3), 190-204.

# Appendix 1: Numbers of messages in each thread for 1994, 1995, and 1996.

This appendix contains the detailed figures for the number of messages per thread for 1994, 1995, and 1996.

In 1994 there were 39 threads with 209 messages. The number of messages in each thread was as follows: 24 19 14 13 12 12 9 9 7 7 7 6 5 5 5 5 5 5 4 4 3 3 3 3 2 2 2 2 2 1 1 1 1 1 1 1 1 1 1 1.

- Mean value of 5.4 messages per thread.
- Median value of 4.

In 1995 there were 45 threads with 535 messages. The number of messages in each thread was as follows: 150 46 41 27 26 18 17 16 16 15 15 12 11 11 10 9 9 9 8 6 5 5 5 5 5 4 4 4 4 3 2 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1.

- Mean value of 11.9 messages per thread.
- Median value of 5.

The figures in this year were quite distorted by some very long threads. If you ignore the longest three threads then the mean drops to 7.1 and the median to 4, very similar to the other years.

In 1996 there were 16 threads with 94 messages. The number of messages in each thread was as follows: 24 13 10 7 7 5 5 4 4 4 3 2 2 2 1 1.

- Mean value of 5.9 messages per thread.
- Median value of 4.

The year 1996 was substantially different from previous years due to the re-structuring of the conferences, a lot of the specific problem solving discussion of TMA 2 in 1995 was moved out into separate conferences, an approach which has continued ever since.

# Appendix 2: Stoplist

This was the stoplist used with the concordance programs and in both versions of the VP.

- a about above according across actually adj after afterwards again against all almost alone along already also although always among amongst an and another any anyhow anyone anything anywhere are arent around as at
- b be became because become becomes becoming been before beforehand begin beginning behind being below beside besides between beyond billion both but by
- c can cant cannot caption co could couldnt
- d did didnt do does doesnt dont down during
- e each eg eight eighty either else elsewhere end ending enough etc even ever every everyone everything everywhere except
- f few fifty first five for former formerly forty found four from further
- g
- h had has hasnt have havent he hed hell hes hence her here heres hereafter hereby herein hereupon hers herself him himself his how however hundred
- i id ill im ive ie if in inc indeed instead into is isnt it its itself
- j
- k
- l last later latter latterly least less let lets like likely ltd
- m made make makes many maybe me meantime meanwhile might million miss more moreover most mostly mr mrs much must my myself
- n namely neither never nevertheless next nine ninety no nobody none nonetheless noone nor not nothing now nowhere
- o of off often on once one ones only onto or other others otherwise our ours ourselves out over overall own
- p per perhaps
- q
- r rather recent recently
- s same seem seemed seeming seems seven seventy several she shed shell shes should shouldnt since six sixty so some somehow someone something sometime sometimes somewhere still stop such
- t taking ten than that thatll thats thatve the their them themselves then thence there thered therell therere theres thereve thereafter thereby therefore therein thereupon these they theyd theyll theyre theyve thirty this those though thousand three through throughout thru thus to together too toward towards trillion twenty two
- u under unless unlike unlikely until up upon us used using
- v very via
- w was wasnt we wed well were weve well were werent what whatll whats whatve whatever when whence whenever where wheres whereafter whereas whereby wherein whereupon wherever whether which while whither who whod wholl whos whoever whole whom whomever whose why will with within without wont would wouldnt

- x
- y yes yet you youd youll youre youve your yours yourself yourselves
- z

# *B882 Additional Electronic Materials Questionnaire*
# *FOR STUDENTS*

**Dear student**

**This survey is being put to a group of B882 students who made considerable use of B882's FirstClass and Web facilities including the experimental 'Uncle Bulgaria' tool.**

**We are trying to see what impact the additional electronic materials have had on your perceptions of the course. We need to know if there are any negative effects and if they can be remedied; we also need to understand the positive effects and how they can be built upon.**

**Your responses will be treated in strict confidence, and will be handled only by a member of the OU's Knowledge Media Institute, who is not connected with the course team.**

**Reports will contain summary information only.**

**We would be grateful if you could provide us with your student ID number, though this is not compulsory.**

**P/I** .....

Simon Masterton, KMi

The Open University
Walton Hall
Milton Keynes
MK7 6AA
Quick Contact: Simon, ++44 (0)1908 655730, S.J.masterton@open.ac.uk, or Simon Masterton on First Class

## *FirstClass questions:*

**This section consists of three questions. The first question is to enable us to understand some of the conditions under which you used First Class. The second question looks at how much you agree or disagree with a number of generalised statements about the FirstClass conferencing system. The third is open-ended and intended to enable you to make any specific comments you have on conferencing as a whole.**

1. Where and how did you access FirstClass? Where? Home, Work
   How? Command line, Off-line reader, On-line reader.

2. Please indicate how well you agree with each of the following statements about FirstClass:

| | Strongly agree | Agree | Disagree | Strongly disagree |
|---|---|---|---|---|
| It is easy to identify the right conference | | | | |
| It is easy to find out what each conference is about | | | | |
| The content of the facilitated discussions is valuable | | | | |
| The content of the free discussion areas is valuable | | | | |
| Support by the special tutors ("water-rats") is valuable | | | | |
| Overall, I was very satisfied with B882 conferencing | | | | |
| I found the e-mail facility very useful | | | | |
| I found the wider MBA conferences very useful. | | | | |

3. Please let us know of any comments you would like to make about the B882 FirstClass facility (continuing on the back of the sheet if necessary):

## *Web questions:*

**This section consists of three questions. The first is for us to get an idea of how many times you have used the web site. The second question looks at how much you agree or disagree with a number of generalised statements about the web site. The third is open-ended and intended to enable you to make any specific comments you have on conferencing as a whole.**

1. How much did you use the B882 Web site? Not at all, A few times, Many times

2. Please rate each of the following components of the B882 Web site for general usefulness. Skip this question if you have not made significant use of the B882 Web site.

| | Didn't use it | Very useful | Useful | Not useful |
|---|---|---|---|---|
| Chef's choice (the short-list of selected sites) | | | | |
| Links to.... | | | | |
| other general creativity sites | | | | |
| related particular course topics | | | | |
| related particular course case-studies | | | | |
| creativity-related discussion areas | | | | |
| "diversions and light relief" | | | | |
| creativity contact information | | | | |
| Web bookshops | | | | |
| Newspapers and journals available on-line | | | | |

3. Please let us know of any comments you would like to make about the B882 Web site (continuing on the back of the sheet if necessary):

## *Uncle Bulgaria questions:*

**This section consists of three questions. The first is to get a taster of how many times you have used the web site. The second question looks at how much you agree or disagree with a number of generalised statements about the Uncle Bulgaria (UB. The third is open-ended and intended to enable you to make any specific comments you have on conferencing as a whole.**

1. How many of Uncle Bulgaria's (UB's) messages did you read? None, Some, Many.

2. Please indicate how well you agree with each of the following statements about Uncle Bulgaria, if you did not read any of his messages please skip this section:

| | Strongly agree | Agree | Disagree | Strongly disagree |
|---|---|---|---|---|
| Uncle Bulgaria's messages... | | | | |
| contained useful information | | | | |
| were relevant to the discussion | | | | |
| answered questions I wanted to ask | | | | |
| provided a different viewpoint | | | | |
| put me off sending my own messages | | | | |
| I would prefer to ask UB questions directly by private e-mail, rather than through the conference. | | | | |
| The name Uncle Bulgaria is a good choice | | | | |
| Uncle Bulgaria should continue to be used on this course | | | | |
| I would like access to all the information UB has. | | | | |

3. Please let us know of any comments you would like to make about the use of the Uncle Bulgaria tool on this course (continuing on the back of the sheet if necessary):

# Appendix 4: Comments from the FirstClass section of the 1997 questionaire (Appendix 3).

[1] "I have just commended the conference structure to my employers as a way of containing the number of e-mails one has to read. Many people copy one-to-one messages to whole conferences unnecessarily, which is an irritant."

[2] "Wigwam was much easier to use + follow"

**Note: WigWam was the offline reader for CoSy.**

[3] "* not sure of the identity of the 'water-rats'

It would be useful if FC personal could hold resumes of either contributers to the offline conferences, or perhaps to a requested set of contributers (like cosy)

The use of two servers can be limiting. it might have been better to use a single server even if that meant a different OS

Backups are intrusive. could the server(1) be mirrored instead"

[4] "The basic idea is correct but it does not match needs. when studying at a traditional university conversations/discussions about the course etc. allow students to explore & understand areas of difficulty. This what first class should be but is not. perhapse small groups matched to tutor groups would help on specific times to discuss so the 'conversation' can be live & broken by days between question & answer. perhaps a working party of some students could operate in different ways to assess if they are better.

Tutors should take more part in the operation & cut into long mono word conversations between two people, i.e.

A - you going to seminar

B - yea

A - have you done the case study

B - yea

etc..."

[5] "There were to many conferences. towards the end of the year there was very little activity, would be better if there was specific topics posted, i.e. mini tutorials or questions from the text with model answers."

[6] "It was useful to keep in touch with other students, and also to feel that support was always available."

[7] "There tends to be a lot of trivia - unfocused dialogue by users in conferences intended to be on course topics. chat areas are the correct place.

Finding a conference is OK once you are familiar with the system but people non computer literate find difficulty. some important course topics are burried as sub-conferences and are thus less obvious. once found you can use the alias to put onto your desktop.

what where the 'water-rat' support tutors - I had not come across these at all.

conference theme is only disclosed by moderators 1st message perhapse a theme banner could be used."

[8] "keep it free of charge!"

**Note: A comment on OU plans to charge for access to the web, this would also affect students coming into FC using the OU's web access rather than direct access.**

[9] "I wish I had more time to participate more fully in the conferences"

[10] "I beleive it needs to be better structured so that it is easier to find worthwhile conferences. It is so easy to get lost in the conferences that I felt had little value & never find the ones that are interesting and useful. Due to the level of 'noise' in the conference I became less inclined to look at the conference. I am particually disappointed by my tutors lack of participation as I find the smaller conferences, based around tutor groups the most worthwhile."

[11] "I never visited the cafe areas etc but I know that they are popular among students. the main problem is users conferencing in inappropriate conferences and being unwilling [unreadable] - hardly something you can police!!"

[12] "useful for summaries from other students

many people fill the screen with type - makes it very hard to read + [three dots meaning 'therefore'] understand."

[13] "I think that the course team created too many conferences.

In particular I did not like the idea of regional conferences. Surely the big advantage of cmc is that location does not matter. ~~The contributions to these conferences were fragmented.~~ [Struckthrough by student]

I found it frustrating that my tutor does not use first class so I could not e-mail him."

**Note: This student was connecting from overseas, a lot of conferencing went on in some of the regional and tutor groupings, making it harder for foreign students to participate, leaving them with only the public discussion conferences. (which had been criticised by others for poor content).**

[14] "Dissapointed by lack of substance in conferences & poor response. Felt I put in more than I got out of it so my interest waned later in the year."

[15] "Its all a balance of time. (see web reply). In principle I feel the facilities are good, but in practice I only used them activly at start of course."

[16] "Overall, I do not consider that I have gained very much for the time expended on this facility.

Frequent breaks in the connection have been an added frustration."

[17] "This system is a major step back from cosy. there was a lot less discussion on the boards than in previous years. several students, I know, refused to use it. the means they did not benefit as much as previous years."

[18] "Main point would be for people to keep to the point of the conference.

too much waffle/chat which should be sent to specific groups of friends - not general conferences.

- information overload."

[19] "too many anoraks in some conferences"

[20] "unfortunatly, too much general 'chit chat' making some of the folders / sub conferences very large & unweildy."

[21] "General facility is excelent, the problem is those students who do not use even the individual regional conferences. look at region 10 for an example. it should somehow be made compulsory to communicate through this conference but not sure how!"

[22] "over complicated structure - regionalisation unhelpful for content discussion."

[23] "I was suprised to find how useful & easy to use it was. I Didnt use COSY at all last year - it was too difficult to get started, & didnt seem to be worth the bother"

[24] "conference structure is messy. subconferences often folders are not visable on desk-top. what about a conference about what is happening inthe B882 conferences with: list, subject, subconferences, new items etc.."

[25] "Access:

I found the wider MBA confrences rather hard to track down. I also have a problem with my software that I cannot read the messages on block - I tried and failed to sort it out. However

generally once I got the hang od managing my desktop, I found access very easy.

Value:

The reason I chose the OU is it is a distance course which I can do to fit in around a full time demanding job and a small demanding son! to suit me best I need to be able to study when I can - any time day or night. The conferencing system is wonderful so that - I am not a good attender of seminars and tutorials because they eat into my time with my son, so the conference provides me with access to other students and ideas. So I am sold on it.

Within this years B820 conferecing I found the TMA conferences, the voluntart sector TMA linked conference and the summaries of most use. I think that for a number of reason:

- Vol sect was well facilitated, soon split into identifiable sub conferences and SMALL (very important I think for flow) about 20 regular participants(?)

- TMA ones provide access to the other students ideas but to really work need to be checked fairly regularly otherwise you lose the thread. I have asked questions of those I didnt understand and got answers and contributed to the debate.

- Summaries is probably most useful in the course term, I use it to give me access to other students interpretation of course material and an addition to my own notes.

I have also used other conferences for work-related questions losley related to the course - mainly vol sec main conf. again very useful. I have also been part of and set up sub conferences arising from residential school.

Downside

? not much. you need to persever to master the system - if you are not getting much out of it, perhaps it dosent seem worth it."

[26] "The B882 conferences are maybe too wide ranging, too much to take in if you have a busy schedual! This is not to say they are not useful, perhapse more focus/facilitation would be useful. Can you analyse who accesses what?"

[27] "Took ages to set up and M.K. helpline was not helpful. London number eventually sorted me out. Students should be warned of costs involved (phone bills) with advice on how to minimise cost. would have liked a regional tutorial, hands on IT day - too little time/too few PC's at the resi sch."

[28] "More instructions on how to use the different FirstClass facilities would be helpful for those who have never used computers befor."

[29] "I found first class a little bit frustrating. it was dominated by several keen people who appeared to use it as an arena to show off. It was full of pointles, meaningless and time wasting chat which for me reduced its usefulness."

# Appendix 5: Comments from the Uncle Bulgaria section of the 1997 questionaire (Appendix 3).

[1] "I havent a clue why Uncle Bulgaria is called Uncle Bulgaria! It is likely as conferences build up that a backlog of similar information will build up and access to this could speed up debates. However if access to information were to be "in block" it would be just like course material and I think loose some of the learning experience of interactive conferencing ie its passive."

[2] good idea - keep it going

[3] I feel that the name had an adverse affect on what the objectives of UB were

[4] Did not find of any great use, and did not fully understand the concept - where was it explained?

**Note: Explanation messages were in the top level welcome conference and all coursework conferences UB was active in.**

[5] was a bit confused what it was all about and had difficulty following the thread of discussion

[6] (I read <u>all</u> I saw, which was not many) [Read All Messages]

I think UB is and interesting experiment and what better course than B882 to experiment on?

It needs an unusual name so it is clear that it is not a student or tutor. Unfortunatly there are always some who do not read the messages explaining what UB is and then post a message asking "who is Uncle Bulgaria?"

[7] useful at times: continue with it

more 'white space' on sum please.

[8] I felt UB helped structure/ensure good quality information in the conference. One of my critism of conferencing is that tutors (as they would in a real tutor group) do not control poor quality student input.

[9] These tended to reflect points raised by other contributors to <u>ask</u> questions to stimulate debate rather than be informative.

[10] [in answer to statement 'I would prefer to ask UB questions directly by private e-mail, rather than through the conference']

Depends on the question - I feel free to use either medium without constraint.

[11] [in answer to statement 'the name Uncle Bulgaria is a good choice']

no feelings either way its just a name to recognise. made me think of parrots!!

[12] Needs developing, perhaps could provide the compliation type answers on individual request?

[13] Once again, perhaps I haven't made enough use of this tool because I failed to spot its potential usefulness.

**Note: The following two quotes were passed on informally by a tutor to me as something students had said during a discussion on the utility of Uncle Bulgaria.**

[14] [The student said that the messages from Uncle Bulgaria implied] you silly people, this topic was discussed ans sorted out ages ago.

[15] some of the contributions are appropriate, others are a tad on the side, but this may be useful as it may be providing a view from another angle

**Note: The following comment came from a tutor in response to a message from Uncle Bulgaria and is included here as it has been used in the analysis.**

[16] Uncle Bulgaria was a Womble, wasn't he? I saw a live descendant on Wimbledon Common today, going through the undergrowth ,putting coke tins in a bag - quite voluntary! What has this to do with B882 - I don't know . Is Uncle Bulgaria a metaphor for something ?- Please tell us.

By the way Uncle B. is not very structured. Could he have some paragraph numbers or something?

Forgive the playful approach

# Appendix 6: Analysis of user types from discussion conferences and relation between number of AA messages read with questions asked.

## Analysis of user types from discussion conferences

Using the history function in FirstClass it is possible to make some analysis of user types and behaviours. Taking the history of all messages posted to the Coursework, TMA1, and TMA2 conferences, there are 7 identified (and overlapping) sets of user. Figure A6.1 shows the sets with the numbers of user identified in each. The sets are as follows:

- **All Users:** all unique user names identified from the 'history' entries of all messages (623).
- **B882 Students:** all users identified as being registered for B882 this year (560).
- **Users Reading AA Messages:** all users identified from the 'history' of the AA messages (430).
- **Students Reading AA Messages:** all users identified as registered on B882 identified from the 'history' of the AA messages (402).
- **All Participants:** all users who contributed at least one message to a conference (117).
- **Students Participating:** all B882 students who contributed at least one message to a conference (105).
- **Students Asking Questions:** the number of students who asked the AA at least one question (40).

From data used to compile each of these sets it has been found that *all* students who have asked the AA a question also participated in the conference (i.e. Students Asking Questions is a subset of Students Participating).

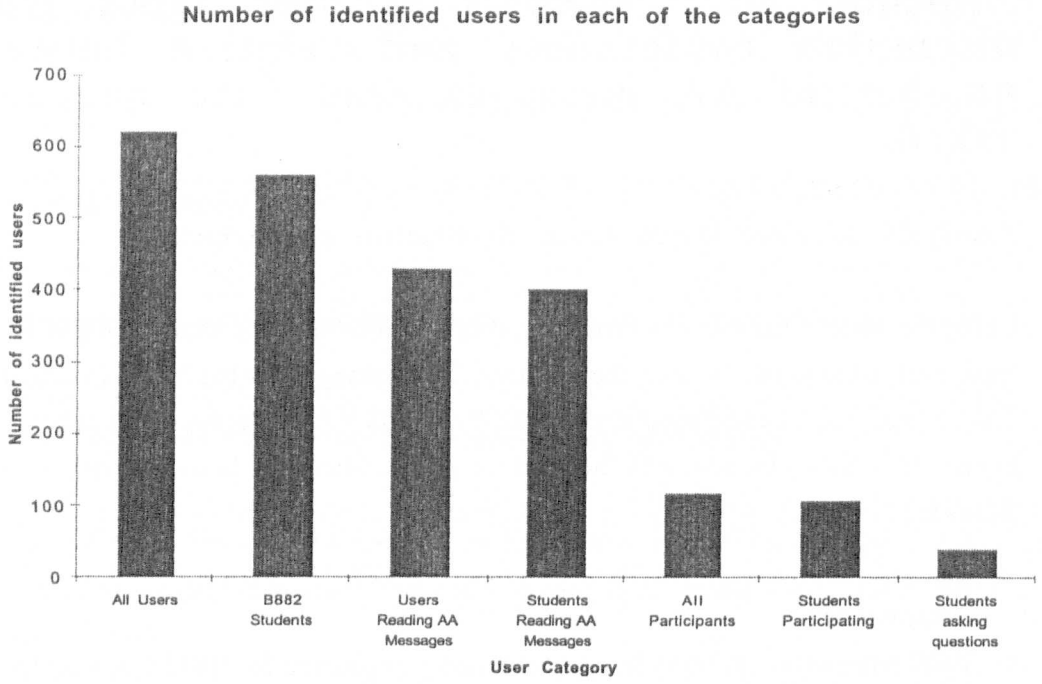


Figure A6.1: The number of students identified in each of the sets.

## Relation between number of AA messages read with questions asked.

Next, I will look at whether the number of AA messages read has any influence on the number of questions the student asks. This can be assessed simply by plotting on a scatter diagram the number of questions asked by students against the number of AA messages read, shown in figure A6.2. Using the Pearson product moment correlation coefficient we get an r value of 0.437 which is a positive correlation with p of 0.01. However if we eliminate our most prolific reader this value drops to 0.277 which is a positive correlation with p of 0.10. This indicates there is some positive correlation between numbers of messages read and numbers of questions asked, but it is not highly significant.

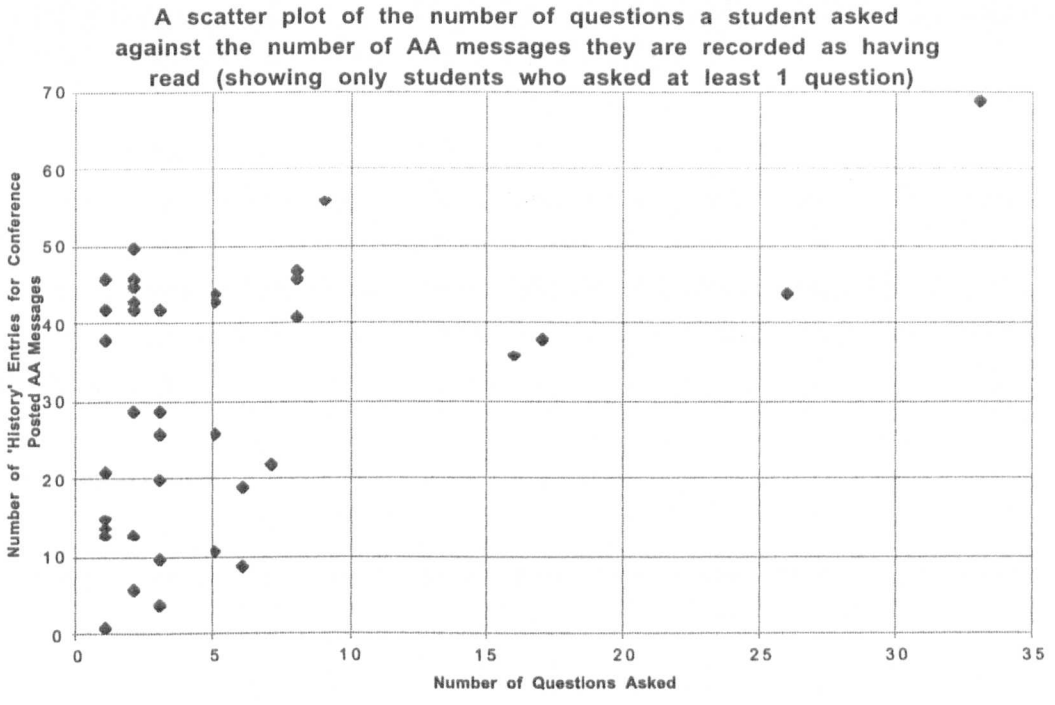


Figure A6.2: A scatter plot of the number of questions asked by each student by the number of AA messages they read.

# Appendix 7: Message contents and thread diagrams from the second study.

### 1. "Reconciling MBTI & KAI Scores"

This thread was just two messages long. The first message was from the student concerned about the relationship between their KAI and MBTI scores. The VP follows up with a reasonably good match with a previous years student who was disturbed about their KAI.

The student posted:

"I have a KAI score of 107 which suggests that I am an innovator. Yet my MBTI profile is ISTJ which would suggest that I had more of the characteristics of an adaptor e.g. practical, logical etc. Can anyone offer an explanation for this?

Alex."

The VP followed up with:

"Disturbed about KAI score

In a previous year one student felt quite disillusioned with their KAI score and felt that either they should become a monk or that all B882 students were doing this course 'in the vain hope of finding at least one organisation which is innovative'.

This was the response of one of their peers:

---

I couldn't agree more. Having scored 126 and working in a German bank !!!! it's no wonder I am having problems with TMA 01 for fear of what I might have to face. However my extension runs out so I SHALL finish it this week. Maybe we should all get together and form a company, we should easily be able to compete with 'the new environment, demanding a greater pace of innovation' all the time.

Plus it might be fun !

---

If you would like to know more about this topic highlight the statement(s) below that you are interested in, go to the message menu, select the sub-menu Reply Special and then the option Reply Sender. When the message has been created, send it and you should recieve your reply shortly.

Why shouldn't we be disturbed with the results?

Is there a link between MBTI/KAI and Star Sign?

---

Please give feedback on this message. Highlight the two statements below and select Reply Sender. In the message generated place the character X in the most appropriate box. Feel free to include any additional comments.

Relevance to Discussion: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ].

Useful for You: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ]."

## 2. "Set"

This is quite a short thread. The students were discussing personal style and development, and the integration of new learning. The VP matched a thread in which a student talks about how their "tutor describes it [the TMA] as a triangle of personal style, organisational style, and development area." The following is the thread diagram for the "Set" thread.

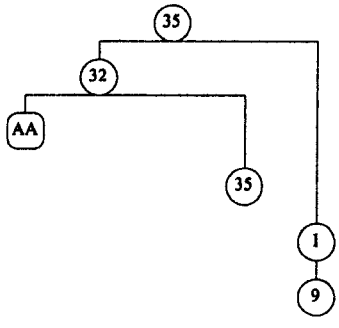


Figure A7.1: Thread diagram for "Set"

### 3. "Personal characteristics and creativity"

Again this is quite a short thread and involves a little discussion between some students and a couple of tutors about different personality types and characteristics and how they influence creativity. The VP in this case retrieves a thread on interpreting MBTI scores. The message from the VP proves so popular that one of its questions reaches the 'general interest' threshold.

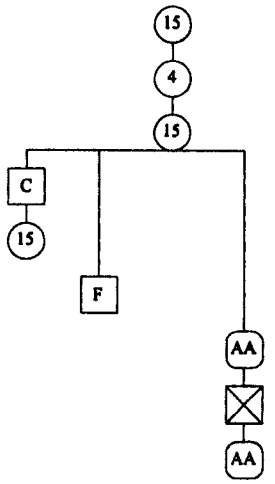


Figure A7.2: Thread diagram for "Personal characteristics and creativity".

## 4. "Creativity or Relativity"

This thread starts off with a discussion about creativity, who are the most creative people in the work force, where to find them, and how to encourage and manage them. The VP then matches a thread which deals with KAI score distribution, what the distributions are in the course, and in industry in general, how creative managers feel in their environments, and the influences they can have. There was a popular question about whether low score on the KAI test meant that you were not creative. However this did not reach the required 'common interest' threshold.

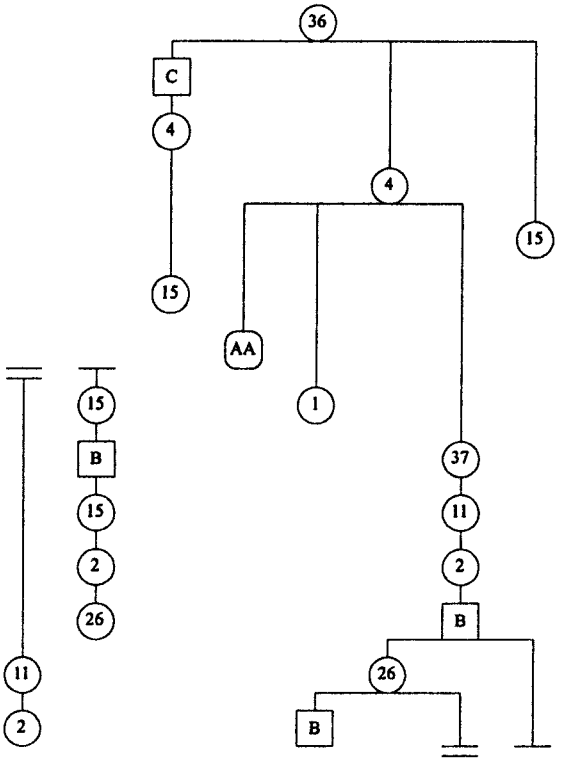


Figure A7.3: Thread diagram for "Creativity or Relativity".

## 5. "Entrepreneurial, Champion and Creative leadership traits..."

This thread proved to be quite a detailed discussion about the traits of different people fulfilling the roles of entrepreneur, champion, and creative leader. Part of the discussion was based on past experiences. The VP matched a thread about how self-employed students could approach the TMAs by pulling details from their past experiences. Not a particularly good match, although one of the questions, on how to approach the TMA, provoked some interest. A feature of question-asking discovered here was that one student kept sending messages with blank bodies. This seems to have been because they were using an earlier version of the FirstClass client than the expected standard. The VP did nothing to respond to these messages (but it certainly should have done).

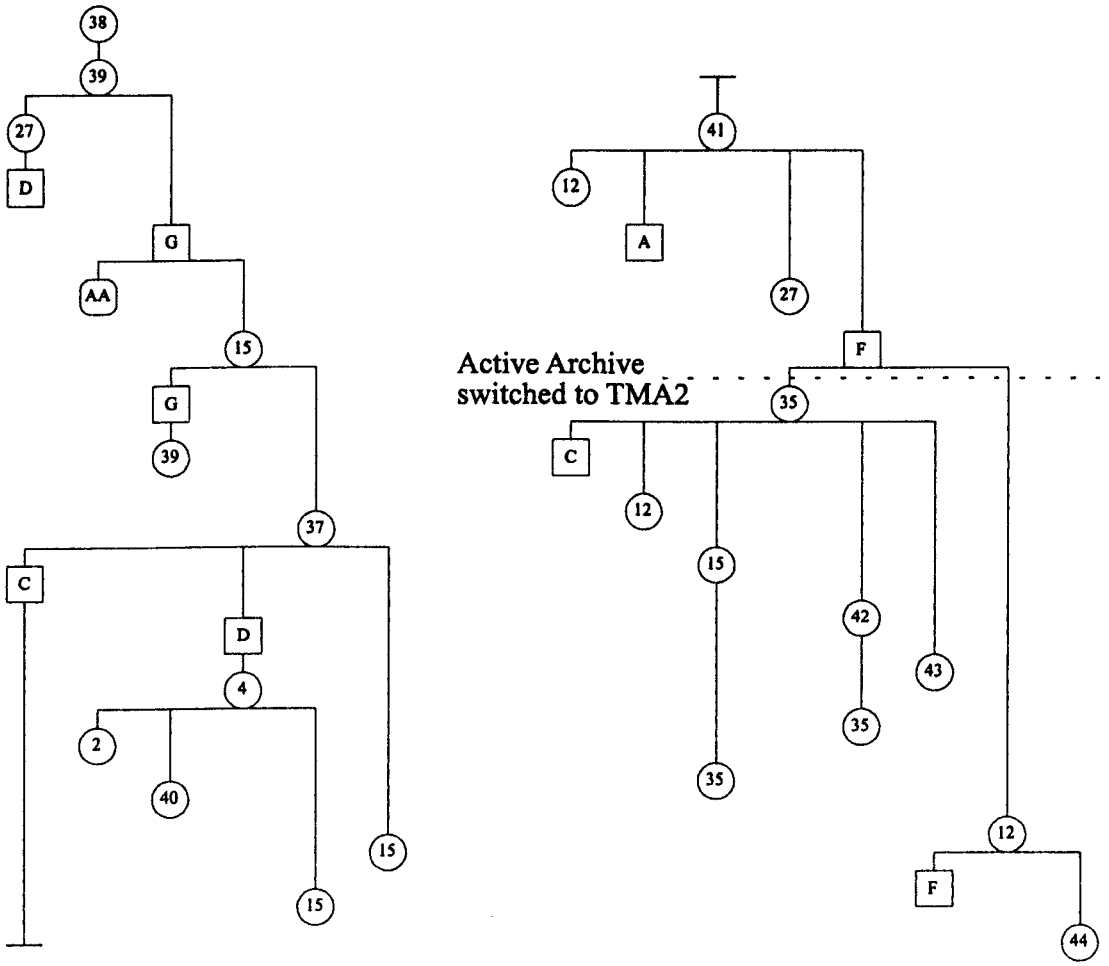


Figure A7.4: Thread diagram for "Entrepreneurial, Champion and Creative leadership traits..."

**5. "Reader, Readings, Readings 2, "**

This was the shortest thread that the VP interacted with during the second TMA of the second study. The discussion revolves around finding where the second readings were in the course materials, and then on the contents of one specific reading. Part of the discussion was on how to continually maintain creativity. The first message from the VP was about creativity warm-ups and ways of stimulating the creative process. The conversation then digressed into opinions on, and recomendations of, other management texts and the problems they show and the way managers faced them. The VP matched a thread about how to treat working though the second block as a problem-solving exercise. There was enough interest in both VP messages that in each case one question reached 'common interest' level and was posted to the conference.

One interesting occurance in this thread is the 3rd message. The student who wrote it used the off-line reader and sent the message at the beginning of the day. However they did not logon, so the system did not update, untill the evening. In the mean time two other messages had been posted. So although it was the 3rd message sent it was the 5th recieved, and the AA processes messages immediatly when they are recieved. Any user looking at this discussion after that day would see the messages ordered in 'sent' order. This did not impact on the processessing of this thread, and is the only exampe of this phenomena observed in a thread the Active Archive contributed to.

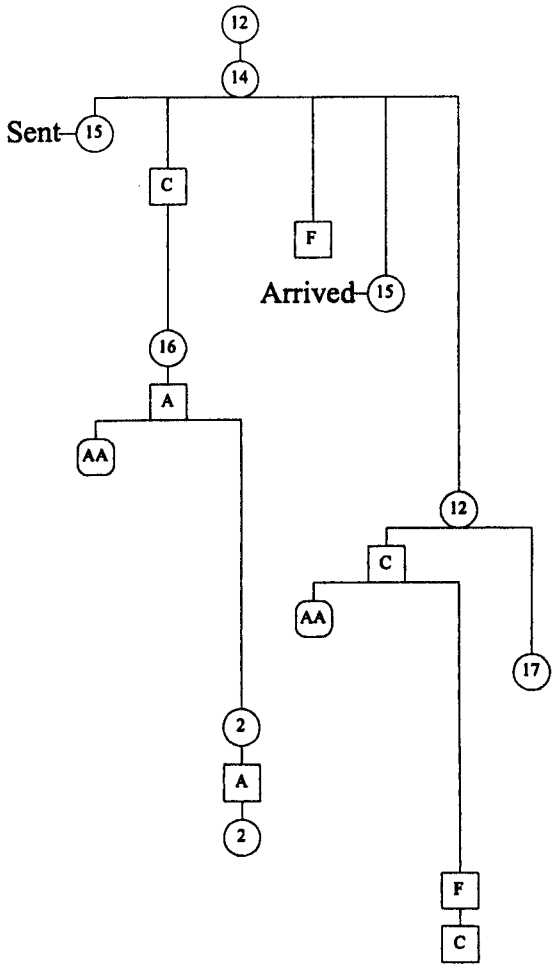


Figure A7.5: Thread diagram for "Reader, Readings, Readings 2, "

# Appendix 8: Stimulated discussion from the 'Team Building' thread.

This appendix contains, in the order which they were posted, the text of the messages from the stimulated discussion of the 'Team Building' thread from TMA2 of the second study. Names have been removed where appropriate.

**16:**

"I think that after the obvious innovation that Semler talks about in Maverick, the most important issue he discusses is reponse to change. The changing environment has caused several major direction changes at Semco. Semler achieved sucess by getting team support for his ideas. The "Satellite" programme was entirely voluntary but would not have worked without the cooperation of the workers involved. As the book progresses Semler steps back from the business and lets the team carry the load."

**AA:**

"Creating a vision

Creating a vision as a problem:

---

TITLE: Creating a Vision as a Problem

My problem for TMA2 is to define what my department's vision should be given that we are just going through major organisational change. I should state that my department is a Public Health Laboratory providing pathology and clinical services to NHS Trusts, GPs and others. Our main role, however, is to protect the population form infectious disease. Does anyone else have a similar issue?

---

A second student had a similar problem:

---

COMMENT: Who's Vision is it anyway

I guess I have a similar problem in that we are a subsidiary of a company that manufactures the products we sell. The parents focus is from a manufacturing viewpoint and toward their local market. What vision can we have of our own, when we are so constrained ie can you lead a market if you haven't control of the products you sell.

The vision has to fit and be realistic. Doesn't mean you can't have longer term changes in view but the vision has to fit now if it's going to be credible.

I guess you have to incorporate what you do now, with what you could offer to your 'customers' in the future. You also have to find out what everyone else thinks assuming you want it publicly adopted rather than for your personal use.

My organisation has a mission statement and set of objectives which are used locally. It incorporates our aims to inlicence products and move further into our market place. I would be happier if we published it to our customers and to our parent company.

My organisation is a pharmaceutical company.

---

If you would like to know more about this topic highlight the statement(s) below that you are interested in, go to the message menu, select the sub-menu Reply Special and then the option Reply Sender. When the message has been created, send it and you should recieve your reply shortly.

Do visions have to be consistant?

Do you mean Vision or Mission?

Visions must be consistant!

Please give feedback on this message. Highlight the two statements below and select Reply Sender. In the message generated place the character X in the most appropriate box. Feel free to include any additional comments.

Relevance to Discussion: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ].

Useful for You: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ]."

**22:**

"I have a similar problem as the small European part of a large but very US focused pharmaceutical company.

From our perspective the US want to control everything - so what is new - seem jealous of our abilities and have no interest in supporting the growth of a successful European organisation, in fact seem to us to be positively trying to stop it.

From their perspective, we are probably upstarts from these strange little European countries who mess up their perfect organisation and keep complaining that they 'do not consider the world outside the USA might have an opinion'! If it is OK in the USA it must be OK in little Europe?!!

There seems to little understanding of the economics of pharmaceuticals today, that if Europe does not contribute to sales, the NPV of most projects is very weak if not negative.

I have the task of trying to persuade them to include European staff in development teams and consider the ex-US markets and needs from the start. This is my TMA 02 planned subject but at present it seems insurmountable.

Has anyone been through this themselves and can offer any unusual suggestions about how to kickstart this process, especially with teams who are very reluctant."

**AA:**

"Stakeholder analysis

Stakeholder analysis is an often used tool, here are some experiences:

---

TITLE: Stakeholders in part (d)

Any one out there using stakeholder analysis in their problem? If so how about sharing your views on the process here so that we can network and get lots marks for part (d)?

I used it with my team last week and found it very good - a few observations to kick this off:

1) You need to prepare well. I used the definitions of the 7 stakeholder types given in the block 2 technique as handouts.

2) You need to stick to the process. A few times it looked like the session was drying up, and we had to take short breaks.

3) It is time consuming - it took us 5 hours x 5 people = 35 manhours

4) It did produce some very useful results which could be easily turned into actions.

---

COMMENT: We used stakeholder analysis as part of the problem exploration stage.

We were trying to define a vision for our departments future. It took us about 15 minutes to do. We took the axes of interest/power. I found it very worthwhile.

---

COMMENT: STAKEHOLDERS

we use it quite often . It is covered in B887 (public sector ) quite well - one of the points of that course is that managers in that sector have more and more diffuse stakeholders to deal with . I was surprised at the time taken - maybe our ' quick and dirty ' version needs to be re-examined . How many stakeholder groups did you identify?

---

COMMENT:

We identified 85 stakeholders - a mixture of individuals and groups. We then 'matrixed' into the Antagonisic, Problematic, Advocate and Low Priority categories. (The block 2 technique really helps pull them out)

Once this stage was complete we looked at potential coalitions between problematic/low priority and antagonistic stakeholders.

Four potentially powerfull coalitions were identified.

The process then concentrated on:

1. strategy for preventing/restricting the coalitions
2. developing the support of the advocates
3. how to get the 'problematics' on board.

This then developed a communications strategy based on stakeholder position. Eight elements/styles of communication were identified.Individuals from the team were assigned as 'owners' of each of the communications elements.

I picked up the task of going to Paris, Madrid, Milan and Frankfurt to talk to the stakeholders, I know its tough and a sarifice, but someone has to do it and there are marks at 'stake' here....

---

Please give feedback on this message. Highlight the two statements below and select Reply Sender. In the message generated place the character X in the most appropriate box. Feel free to include any additional comments.

Relevance to Discussion: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ].

Useful for You: Excellent[ ], Good[ ], Fair[ ], Poor[ ], Not at all[ ]."

**24:**

"I am also thinking of looking at team building. We are also going through period of change, and at the residential we also came up with the thought that I had to try and make them see the situation as an opportunity not a threat!! Not easy at the moment as I work in a Trust which merged with another two years ago, and now may, or may not, de-merge. However we will certainly merge with another Trust. Although at the moment it is out for public consulation and the result won't be known until September, unless they delay it for a second time!!

Sorry if I sound cynical but I work for the NHS and it has been a bad day!!"

**20:**

"student 22 writes:

Has anyone been through this themselves and can offer any unusual suggestions about how to kickstart this process, especially with teams who are very reluctant.

Try 'skunkworks' by Tom Peters - mentioned in Maidique or one of the readings. If you can build up sufficient momentum in a small committed local team, then acquire a local executive sponsor, and if you can get him / her hooked by proving the concept, provide a useful deliverable or whatever, he/she can then promote it in the US. I work in such a setup, and I've seen it work. Our quality management system was born in the UK and Netherlands, and after many years the US bods have finally bought in and it is now a corporate standard.

You may find that the US are not dismissive of local expertise, developments etc, in my experience it is frequently that they never get to hear of local developments, and assume none are forthcoming. I suggest getting a suitable executive sponsor is a major, useful asset - so get networking! This could be of immense benefit personally, and indirectly for your local operation. Check out you stakeholders - cultivate those you can influence, by developing a proof of concept, a sample of output etc.

Hope this helps."

# Appendix 9: Further discussion on the impact of the VP on discussions

This appendix contains some further discussion on the impact of the Virtual Participant on discussions. Some of the contents is speculative and the rest provides pointers towards ways of exploring this issue further. A more detailed study into the relationship between the depth and amount of branching still seems appropriate. Personally I believe that both the depth and amount of branching are directly dependent on content and context but this has yet to be demonstrated. Later in this Appendix I will present a brief example from a wholly different domain of online discussion to illustrate the difference the type of subject can make.

**Evaluating Interactions**

Chapter 7 presented an analysis of how many replies a message is likely to receive and compares messages from the Virtual Participant during the second study with those from the students. The conclusion reached was that VP messages cause more branching, but do not have a significant effect on the number of replies a message might expect to get, and therefore by inference no significant effect on the structure of the discussions.

When approaching this question one of the first options to explore is depth of the thread versus the number of messages posted to that thread. The problem here, as with the evaluation in chapter 7, is picking a reasonable sample of threads as in all cases it is not possible to split the students between a test group and a control group. One approach to this is to use threads from the third TMA as a control. The main assumption here is that there were no changes over time in the discussions (which I can not say is definitely true). Possible changes in discussion format may be due to the less committed students dropping out, the change in topic and the necessary level of discussion, and greater experience of FirstClass. However, this still presents a more realistic measure of the same group of students than a previous year's sample.

Taking the depth versus number of messages approach the next step is to create a scatter chart against these axes, or to use a table with the number of threads satisfying a certain combination being indicated. This is shown in table A8.1 below. This table has all the threads from TMAs 1, 2 and 3, apart from the threads of length 1, and one outlying thread of length 31 and depth 11.

| | Length of thread | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Depth** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 2 | | 2 | 2 | 1 | | 2 | | | | | | | | | | | | | | | | | | |
| 3 | | | 3 | 3 | | 2 | 1 | | | | | | | | | | | | | | | | | |
| 4 | | | | 3 | 6 | 3 | | | | 1 | | | | | | | | | | | | | | |
| 5 | | | | | 1 | | 1 | 1 | | 1 | | 2 | | | | | | | | | | | | |
| 6 | | | | | | 2 | | | 1 | 2 | 1 | 1 | | 1 | | | | | | | | | | |
| 7 | | | | | | | | | | | 1 | | | | | | | | | | | | | 1 |
| 8 | | | | | | | | | | | | | | | | | | 1 | | | | 1 | | 1 |
| 9 | | | | | | | | | | | | | | | | | | | | | 1 | | | |
| 10 | | | | | | | | | | | | | | | | | | | | | | | | |
| 11 | | | | | | | | | | | | | | | | | | | 1 | | 1 | | | |
| 12 | | | | | | | | | | | | | | | | | | | | 1 | | | | 1 |
| 13 | | | | | | | | | | | | | | | | | | | | | | | | |
| 14 | | | | | | | | | | | | | | | | | | | | | | 1 | | |

Table A8.1: Showing the number of threads for each combination of number of messages, or length of thread, versus depth of thread.

From table A8.1 it is possible to approximately cluster the data into three groups. Group A is up to 8 messages and contains a large number of threads. Group B runs from 9 messages to 14 messages and is our middle set. Group C contains the high number of messages threads. Those in group A are mostly short discussions consisting of one student asking a question and then receiving a number of suggestions or a short discussion with turn taking (hence depth = number of messages). Group C on the other hand are those threads which have become very involved in a deep discussion. Finally those in group B are intermediate threads. Where group C is open ended, and group A can be answered simply, group B are threads where there is a number of differing points of view, or perhaps confusion, which needs to be managed.

**Future Studies**

It is interesting to look at the relationship between the amount of branching of a discussion against the depth of that discussion and how that relates to the content. I personally think this area deserves deeper study. While making these comparisons I made a quick sample of a Usenet news group I expected to have rather deep and convoluted discussions (talk.abortion) to look at the average thread length there. Taking the first 15 threads that day the average length was 17, with a standard deviation

of 9.5 and a median of 17. In general these were much longer threads (c.f. average 9.5, SD 7.1, median 6) than those observed for B882.

It seems appropriate that a more detailed study could be undertaken in this area, to investigate this hypothesis that length and branching factor are to some extent dependent on content and context.

**B882 First Class usage and opinions survey 1998**

Dear Student,
This survey is from Simon Masterton in the Knowledge Media Institute and forms part of my ongoing research into how students make use of First Class and tools which we can provide to assist them. This survey is being put to most B882 students who have used First Class at some point during this year. I am trying to gauge how you use, and what you think of First Class both on B882 and on other MBA courses. I am also evaluating the experimental Active Archive tool which was posting messages to the discussion conferences during TMA's 1 and 2.
The questionnaire is entirely optional, you may complete as much or as little of it as you wish. This work forms part of an ongoing research project. This questionnaire is completely confidential: in the interests of privacy no student will be identified. However, we may use anonymised quotes from this survey in articles and conferences. By completing this questionnaire you agree to this usage.

This questionnaire consists of 3 sections. The first section asks about your use of First Class on the MBA as a whole. The second section focuses on the use of First Class in B882. The third section looks at the Active Archive.

All questions are optional and I am very grateful for the time you take to respond to this questionnaire.

Quick contact:
Simon Masterton on First Class
or call ++44 (0)1908 655730

**Initial background questions:**

Please state your name and your first class ID, i.e. 'Simon Masterton', sjm36

Please give the course codes of all past and current MBA courses where you have used First Class (FC), or any other form of conferencing.

For the next two questions please circle the option that you normally use.

| | | |
|---|---|---|
| Where do you access FC from? | work | home |
| How do you access FC? | First Class Client | First Class Personal |

**Part 1. General Use of First Class.**
This section is intended to get an overall picture of your use of First Class in the MBA program and not just on B882. Please feel free to add any additional comments at any point using the space on the back.

1) Across the MBA program, as a user of First Class would you say you were a heavy; medium; or light user?

2) Across the MBA, on average how often do you connect to the First Class server,

every day, twice a week, once a week, less often

3) Over the period of your MBA has there been any times where your use of FC has markedly increased or decreased in any way? Yes/No - if yes, please state why.

4) Across the MBA, which facilities do you use regularly - circle all that apply? e-mail-other students, e-mail-tutor, conferencing, private text chat facility, other - please specify

5) Across your MBA use of First Class, please list the courses where you felt using FC helped you to achieve a better mark.

6) Of the courses listed in 5 above, in which did you most value the FC conferencing. Please state why.

7) Across your MBA use of First Class, please list the codes of any courses where you felt that conferencing was superfluous, or did not live up to expectations. Please state why.

8) Please use this space, and continue on the back if necessary, to make any additional comments about any aspect of conferencing on the MBA program.

**Part 2. First Class and B882.**
This section is intended to get an overall picture of your use of First Class only on B882. Please feel free to add any additional comments at any point using the space on the back.

9) On B882, in comparison to the other MBA courses, would you say you use First Class
Less[ ], Same[ ], More[ ].

3) Over the period of B882 has there been any times where your use of FC has markedly increased or decreased in any way? Yes/No - if yes, please state why.

11) Of the B882 conference areas, please rank them in order of importance, staring with 1 as the most important, until all are marked:
coursework[ ], archive[ ], computing[ ], common room[ ], locals and specials[ ], other[ ] - please specify.

12) Are you happy with the B882 conference structure this year, Yes/No
If No - how would you suggest it was changed ?

13) Please state how useful conferencing has been for you:
Very useful[ ], useful[ ], not useful[ ], not at all useful[ ]

14) How do you feel using the B882 Conferences has helped you in your study of the course?

15) How has your involvement in conferencing, be it contributing, browsing, or both, informed and enhanced your performance in your TMA's ?

16) How easy is it to find the right conference?
Very easy[ ], easy[ ], difficult[ ], very difficult[ ]

17) How easy is it to find out what each conference is about?
Very difficult[ ], difficult[ ], easy[ ], Very easy[ ]

18) How valuable is the content of the facilitated discussions?
Very valuable[ ], valuable[ ], not valuable[ ], not valuable at all[ ]

19) How valuable is the content of the free discussion areas?
Not valuable at all[ ], not valuable[ ], valuable[ ], very valuable[ ].

20) How satisfied were you with the B882 conferencing?
Very satisfied[ ], satisfied[ ], dissatisfied[ ], very dissatisfied[ ].

21) How useful did you find the e-mail facility?
Not useful at all[ ], not useful[ ], useful[ ], very useful[ ]

22) How useful did you find the wider MBA conferences?
Very useful[ ], useful[ ], not useful[ ], not useful at all[ ]

23) Please use this space, and continue on the back if necessary, to make any additional comments about any aspect of conferencing on B882.

**Part 3. The Active Archive.**
This section is intended to get an overall picture of your use of the Active Archive for B882. Please feel free to add any additional comments at any time.

24) Have you made any use of any of the archived conferences? Yes/No

If you answered yes, please give more details about what you used. If you answered no, please say why.

25) Please say how the archive conferences could be improved.

26) When seeing a message from the Active Archive posted to a discussion conference would you choose to read
every one[ ], only those on topics you were following[ ], none of them[ ], other please specify

27) Of all the messages posted by the Active Archive, how many would you say you read?
1-3, 4-6, 7-9, 10 or more, none

28) How many questions have you asked the Active Archive over the length of B882?
1, 2-4, 5-7, 8 or more, none

29) Did you find the responses, by the Active Archive, to your questions any use? Yes/No

If not please could you suggest any way in which they could be improved.

30) How often did you give feedback on messages from the Active Archive?
Always[ ], Often[ ], Sometimes[ ], Never[ ].

31) If you did not give feedback, why was this?
Messages always relevant[ ], Too much effort[ ], No benefit to me[ ], Other please specify.

32) How relevant were the individual messages from the Active Archive to the discussions?
Very relevant[ ], relevant[ ], not relevant[ ], not relevant at all[ ]

33) How relevant were the individual messages from the Active Archive?
Very relevant[ ], relevant[ ], not relevant[ ], not relevant at all[ ]

34) Do you feel that you learned anything from either the active archive messages posted to the conference or from any questions you asked it? - please specify/give details?

35) How has the Active Archive has informed and enhanced your performance in your TMA's ?

36) The Active Archive is not available in TMA 3 or 4, although all the archived messages are. Would it be useful if it was continued in these later conferences?

37) Currently the Active Archive is only being used on B882. Do you think it would have proven useful on any other MBA courses you have done? Yes/No

**Please rate how you feel about the following statements:**

38) The name 'Active Archive' is a good choice.
Agree strongly[ ], agree[ ], disagree[ ], disagree strongly[ ].

39) The active archive should continue to be used on this course
Disagree strongly[ ], disagree[ ], agree[ ], agree strongly[ ]

40) I would rather the active archive posted messages directly to those participating in the discussion and not to the conference.
Agree strongly[ ], agree[ ], disagree[ ], disagree strongly[ ].

41) The Active Archive put me off participating in the conference
Disagree strongly[ ], disagree[ ], agree[ ], agree strongly[ ]

42) The Active Archive reduced the amount of discussion
Agree strongly[ ], agree[ ], disagree[ ], disagree strongly[ ].

43) The Active Archive raised useful questions
Disagree strongly[ ], disagree[ ], agree[ ], agree strongly[ ]

44) The Active Archive messages were relevant to the current debate
Agree strongly[ ], agree[ ], disagree[ ], disagree strongly[ ].

**For the next few questions you will be asked to give a few words to describe your feelings about different aspects of the Active Archive. What follows is some examples of words which you might find appropriate, but you are welcome to use any that you might prefer.**

Interesting, helpful, unhelpful, patronising, assisting, distracting, stimulating, exciting, unclear, facilitating, conversation stopping, related, should be used more often, right, pleasant, on course, off topic, content, should not be used, relaxed, faceless, showing off, ambiguous, obscuring, wrong, cryptic, esoteric, plain, ineffective, thread killer, different, same again, different.

45) Please write down a few words to describe the name 'Active Archive'

46) Please write down a few words to describe how you feel when you see a message from the Active Archive.

47) Please write down a few words to describe the content of messages from the Active Archive.

48) Please write down a few words to describe the your overall opinion of the Active Archive.

49) Please write down a few words to describe the your overall opinion of First Class on B882.

50) Please use this space, and continue on the back if necessary, to make any additional comments about any aspect of the Active Archive.

# Appendix 11: Full data from the second survey

This appendix contains important raw data from the *second* survey, and related data from the *first* study, on the student opinions on B882 conferencing.

### Student opinions on B882 conferencing

This section contains the raw data for questions 16 to 22 from the second study and the statements contained in the second part of the first study. The first two tables are from the second study and contain first the raw values and then the percentages. On these scales 1 corresponds to very easy and 4 corresponds to very hard.

| Values for the *second* study | Very Easy(1) | Easy(2) | Hard(3) | Very Hard(4) | T | Ave |
|---|---|---|---|---|---|---|
| How easy is it to find the right conference | 15 | 69 | 31 | 1 | 116 | 2.16 |
| How easy is it to find out what each conference is about | 10 | 68 | 31 | 6 | 115 | 2.29 |
| How valuable is the content of the facilitated discussions | 5 | 56 | 34 | 10 | 105 | 2.47 |
| How valuable is the content of the free discussion areas | 3 | 34 | 56 | 16 | 109 | 2.78 |
| How satisfied were you with the B882 conferencing | 7 | 65 | 29 | 7 | 108 | 2.33 |
| How useful did you find the e-mail facility | 28 | 63 | 16 | 7 | 114 | 2.02 |
| How useful did you find the wider MBA conferences | 0 | 69 | 29 | 12 | 110 | 2.48 |

Table A10.1a: Raw data for student opinion from the *second* study

| Percentages for the *second* study | Very Easy(1) | Easy(2) | + | Hard(3) | Very Hard(4) | - |
|---|---|---|---|---|---|---|
| How easy is it to find the right conference | 13 | 59 | 72 | 27 | 0 | 27 |
| How easy is it to find out what each conference is about | 9 | 59 | 68 | 27 | 5 | 32 |
| How valuable is the content of the facilitated discussions | 5 | 53 | 58 | 32 | 10 | 42 |
| How valuable is the content of the free discussion areas | 3 | 31 | 34 | 51 | 15 | 66 |
| How satisfied were you with the B882 conferencing | 6 | 60 | 66 | 27 | 6 | 33 |
| How useful did you find the e-mail facility | 25 | 55 | 80 | 14 | 6 | 20 |
| How useful did you find the wider MBA conferences | 0 | 63 | 63 | 26 | 11 | 37 |

Table A10.1b: Percentage data for student opinion from the *second* study

**And from the first study the related data was:**

| Values for the *first* study | Very Easy(1) | Easy(2) | Hard(3) | Very Hard(4) | T | Ave |
|---|---|---|---|---|---|---|
| It is easy to identify the right conference | 9 | 20 | 8 | 2 | 39 | 2.08 |
| It is easy to find out what each conference is about | 4 | 23 | 12 | 1 | 40 | 2.25 |
| The content of the facilitated discussions is valuable | 5 | 25 | 10 | 0 | 40 | 2.13 |
| The content of the free discussion areas is valuable | 1 | 21 | 17 | 1 | 40 | 2.45 |
| Support by the special tutors ("water-rats") is valuable | 10 | 25 | 4 | 0 | 39 | 1.85 |
| Overall, I was very satisfied with B882 conferencing | 6 | 20 | 11 | 2 | 39 | 2.23 |
| I found the e-mail facility very useful | 14 | 17 | 7 | 1 | 39 | 1.87 |
| I found the wider MBA conferences very useful | 8 | 17 | 13 | 0 | 38 | 2.13 |

Table A10.2a: Raw data for student opinion from the *first* study

| Percentages for the *first* study. | Very Easy(1) | Easy(2) | + | Hard(3) | Very Hard(4) | - |
|---|---|---|---|---|---|---|
| It is easy to identify the right conference | 23 | 51 | 74 | 21 | 5 | 26 |
| It is easy to find out what each conference is about | 10 | 58 | 68 | 30 | 3 | 33 |
| The content of the facilitated discussions is valuable | 13 | 63 | 75 | 25 | 0 | 25 |
| The content of the free discussion areas is valuable | 3 | 53 | 55 | 43 | 3 | 45 |
| Support by the special tutors ("water-rats") is valuable | 26 | 64 | 90 | 10 | 0 | 10 |
| Overall, I was very satisfied with B882 conferencing | 15 | 51 | 67 | 28 | 5 | 33 |
| I found the e-mail facility very useful | 36 | 44 | 79 | 18 | 3 | 21 |
| I found the wider MBA conferences very useful | 21 | 45 | 66 | 34 | 0 | 34 |

Table A10.2b: Percentage data for student opinion from the *first* study

# Appendix 12: Highlights of student comments in the 1998 questionaire

The majority of comments from section 3 (and the last question of section 2) of the questionaire are included here. For the sake of clarity those which are *not* relevant to the Active Archive have been removed. In each case, the question which students are commenting on is presented first, then the comments themselves.

23) Please use this space, and continue on the back if necessary, to make any additional comments about any aspect of conferencing on B882.

---

[23-1] Conferencing activities need to be better supported for users who are new to this tool

[23-2] This media could be more widly used but would probably have to become compulsory.

[23-3] The only real problem with FC is selection of relevant data. There is so much to review!! - similar problem as all systems & internet. Additional frustration caused by time constraints.

[23-4] If the content is poor then so will the archive!!

[23-5] Conferencing has been dominated by a few people

[23-6] Reams and reams of opinions of people i've never met. I don't have time nor interest in finding it to participate.

[23-7] The conferencing has far too many contributors saying the same thing to make it worth while trying to find the 'nuggets of gold'.

[23-8] The conferences should be moderated to keep people on track.

[23-9] some way of encouraging otheres to contribute more. instead of browsing, needs to be found.

[23-10] The wider the contributions the richer the learning.

[23-11] It would REALLY REALLY help if some indication of the type of content in a contribution could be given before having to go in

and read an email. Many emails are of the "I agree", "My books haven't arrived too", "Thanks to everyone at UMIST for great week" type and it is such a waste of time reading them. If either;

1) contributors self-ranked them in order of relevence;

2) some kind of relevence algorithm could be developed; or

3) readers were asked to rank the email after reading it on a scale

from 'waste of time' to 'must read' and average relevance scores were derrived AND these scores appeared beside unopened emails this would really help to wade though the ever widening swamp.

24) Have you made any use of any of the archived conferences? Yes/No
If you answered yes, please give more details about what you used. If you answered no, please say why.

---

[24-1] Unclear about the purpose of the archives

[24-2] Just looked in briefly to see what ideas were there

[24-3] Look up people who had been in a similar situation to my self and how they approached the TMA - also how other people interpret questions and structure answers

[24-4] Visited TMA archive prior to completing my assignment

[24-5] To obtain additional information on specific subjects quickly

[24-6] Did not understand what it was from the start

[24-7] I cannot remember specificically, but it was a summary of the previous years comments.

[24-8] I tried to get further materials on several conference for TMA01 - Followed instructions but couldn't get a reply. I was busy so I gave up + haven't tried since.

[24-9] asked for info on specific topic.

[24-10] Helped answer some of the questions I was asking.

[24-11] used past tutor comments and read all questions

[24-12] I have tracked the responses and the tutors opinions

[24-13] Use of Active Archive to gain pointers to topic areas and discussion threads that I am unable to locate by other means

[24-14] Assumed relevant ones would be posted to discussion conferences.

[24-15] Too enigmatic(!!) when I started so I gave up.

[24-16] Too much volume anyway. Really need a proper search/query engine

[24-17] Too much work, no need felt; if necessary, it was easier to ask than to wade through archives. FC personal is not really geared to useing it.

25) Please say how the archive conferences could be improved.

---

[25-1] Possible comments from tutors a validity of comments also some comments appear to be only vaguely linked.

[25-2] more selections - better 'user interface'

[25-3] Needs to be more interactive, seemed to be a summary of past conferences with insufficent opportunity to discuss content - "Too Dry".

[25-4] Shorter - more precise comments

[25-5] Sometimes the relevant to the message thread was a little obscure.

[25-6] They seeemed dead - like last weeks newspapers

[25-7] Take them out of threads if they do not relate

[25-8] They could be given their own conference. Even a conference per archive

[25-9] group more "old" messages around subject

[25-10] They should be better described, such as a summary in one file and then anyone can decide what should be downloaded.

[25-11] Proper query/search engine

[25-12] Allowing readers to access archive directly.

[25-13] Put on the web, including search facility? - maybe by moderators referring to it based on topic when something is discussed is the main conference. Or place a good structure in one mail, sort of table of contents, so that it is easier to search.

29) Did you find the responses, by the Active Archive, to your questions any use? Yes/No

29a) If not please could you suggest any way in which they could be improved.

---

[29-1] More useful that most contributions.

[29-2] some answers too general or not very specific don't know original context of message

[29-3] Problem was that many of the messages posted in the conferences were off-topic and at some stage, dozens were generated filling up the whole conference. The response I got was 'more or less' on topic and usefull, but the contents itself was not of such a standard that I got a lot out of it that I could not have found in another way, like conferencing with tutors/students or reading books.

[29-4] Maybe a search facility like Altavista in combination with a structure or table of contents to access the archive would be more effective, although I like the idea of the active archive. The problem is the amount of information available (as always!): it is very time-consuming to find what you are looking for. The active archive does not really allow to browse to related subjects, so if the reply is off-topic, you can not search further.

31) If you did not give feedback, why was this?

Messages always relevant[ ], Too much effort[ ], No benefit to me[ ], Other please specify.

---

[31-1] Not clear instructions

[31-2] Was not sure how to, it was probably very easy.

[31-3] I tried but that macheine did not permit it.

[31-4] as with all conferencing, the postings will be of benefit to someone, somewhere.

[31-5] Talking to an automation thats really sad

[31-6] I didn't feel I could contribute

[31-7] no feedback re. recipt of messages left me doubting that 'feedback' was actually used.

[31-8] usually dumped offline, didn't realise could reply.

[31-9] I'm paying for this call

[31-10] too impersonal

34) Do you feel that you learned anything from either the Active Archive messages posted to the conference or from any questions you asked it? - please specify/give details?

---

[34-1] Helpful as expressed previous students ideas of TMA structures

[34-2] Gained different viewpoints

[34-3] Yes, primarily to do with the structure of TMA's

[34-4] Yes: e.g.interpretation of KAI score.

[34-5] Yes - gave some leeds for TMA subjects

[34-6] yes for AA, no for others [refuring to other students contributions].

[34-7] Have helped to broaden understanding of various topics without wading through drivel

[34-8] minimal - only really a feel for topics others were including in their TMAs.

[34-9] enabled to balance viewpoint and look @ topic from a different perspective

[34-10] no - dismissed it - didn't light a fire!

[34-11] Very relevant messages - seriously impressive database management. [however, when it came to learning from it] not really - the cleverness of its being there means it has undue weight attached to it, whereas in fact its content is only the same as that of the current discussions

[34-12] Not really learned, although it was nice to read it - as it was a transcript from a student question, it is interesting to se that others have the same type of questions.

37) Currently the Active Archive is only being used on B882. Do you think it would have proven useful on any other MBA courses you have done? yes/no

---

[37-1] It would probably be even more benificial on a more 'structured' course

[37-2] difficult to see how you could weed the stuff you've got on courses where it would be useful because the same case studies are being used year after year, e.g. B820

45) Please write down a few words to describe the name 'Active Archive'

---

Archive sounds too mechanical, beauracratic, and involving effort

Stimulating Confusing

No major experience => ambivilent

living past, gone but still here, history today

| | |
|---|---|
| Assisting. | faceless, oxymoron |
| assisting, related | interesting, useful |
| More interesting - motivative - rareties "old" | deceptable |
| Ambigious | related, conversation stopping, distracting, helpful |
| catchy, trendy, brand | confusing, technical jargon, not user friendly |
| Stimulating, thought provoking. | descriptive, effective |
| Appropriate | Helpful, give more ideas |
| Not very user friendly - how about "kilroy" | facilitating + interesting |
| Providing information form a managed source of data | assisting, stimulating |
| Still need direction as to best use | helpful |
| Useful | asking, facilitating, unclear |
| Assisting, facilitating, relevant | former/old message concerning a ceartain subject, made retrievable |
| assisting | cryptic |
| Helpful, facilitating | interesting |
| Helpful | interesting, stimulating |
| Interesting, helpful | relevant, reflects service provided |
| good sound & image phrase | cryptic, distracting, unclear |
| what? | |

aged, tired, impersonal

obscure

unclear

appropriate/ relevant

too clever by half

Different, helpful

a means of retrieving and posting relevant comments from previous years' students on this course

Useful information resurfaced

cryptic

Interesting, helpful, should be used more often, faceless, different, unknown

irritating, patronising

ok, sometimes overactive archive

sometimes patronising but at times helpful

Useful

46) Please write down a few words to describe how you feel when you see a message from the Active Archive.

---

[46-1] positive - look at it like any other message

[46-2]Usually read it since it has more chance of being a useful contribution.

may be relevant - only scan unless interested

Interested, curious

who has written this?

Intrigued

?not yet experienced

Interested

usually skip them

helpful

pleased, interested

Better check - "Just in Case"

helpful

distracting

may help

on course, related, facilitating

curious

pleased, interested.

Interested

interested

relaxed

curious

interesting

pass, next, F2

intrigued & curious & expecting

interesting, sometimes too much

exciting

same again

curiosity

interested

content

unclear, unsure

conversation, on course, different

faceless, off topic

I don't see them very often [recorded 15, claims 4-6]

indifferent

Interesting, helpful

I wonder what specificially its about

might be useful

move to the next message

I feel unclear of the intention of the sender

Interested, doubtful

perhapse too many instructions on screen with the message

oh god there's more to get through

interested, curious

same again

47) Please write down a few words to describe the content of messages from the Active Archive.

---

| | |
|---|---|
| could be more specific to conference | interesting, mixed |
| Some spelling and gramatical errors annoy me | often useful, somtimes cryptic |
| Variable sometimes ambigious | dead, missing the point, |
| Relevant | helpful & assisting |
| okay | interesting |
| Relevant | related |
| Related, unclear, cryptic | assisting but also "is this all" |
| helpful | helpful, interesting |
| facilitating, faceless, different | related, facilitating |
| a bit long | ineffective, faceless |
| thought provoking | helpful, interesting |
| helpful | havent read enough [claims 7-9, records show 4]. |
| relevant | not that useful |
| to the point | helpful |
| appropriate, guiding | a lot is superflouous |
| Direction, slant | interesting |
| relevant useful | eccentric, off-key, dispassionate, detached |
| helpful, plain | relevant and good |
| useful, interesting, often valuable | |

| | |
|---|---|
| relevant, but historic | off topic (in conferences, not in direct reply) |
| helpful, interesting, meaningful | helpful |

48) Please write down a few words to describe the your overall opinion of the Active Archive.

---

[48-1] good idea - each message should be summarised at start with short sentence of contents

[48-2] Could be better - easier to use - more instructive

[48-3] A good innovation with promise for other courses.

[48-4] distracting but should be used more often

[48-5] a possibly powerful tool

[48-6] Should be used more often

[48-7] Just key infor would be better - delete superflouous stuff.

[48-8] informs current discussion

[48-9] clever, but over the top. Nice to know it can be done, but so what?

[48-10 opportunity for improvement ;-)

| | |
|---|---|
| wonderful, faceless | okay, useful, interesting |
| Extra Info | helpful |
| Useful if I have time | indifferent |
| Innovative | innovative |

| | |
|---|---|
| | Usually ignored it. |
| helpful, assisting | |
| | good initiative |
| confusing as to me | |
| | stimulating |
| limited use | |
| | facilitating, assisting |
| helpful, relevant | |
| | helpful |
| facilitating | |
| | unclear, ambiguous |
| fine | |
| | stimulating, assisting |
| useful/good | |
| | neutral |
| beneficial, valid | |
| | not that useful |
| useful tool in course | |
| | vaguely interesting |
| nitty-gritty | |
| | tries hard |
| assisting, on course | |
| useful, interesting, valuable | not having used it very much, it is difficult to comment |
| OK, sometimes relevant, always interesting | useful, potential, cumbersome |
| useful | useful addition to conference |

49) Please write down a few words to describe the your overall opinion of First Class on B882.

---

| | |
|---|---|
| Probably over used - with so much activity it's easy to miss "useful" messages | Very useful contact with others is great |
| | Good facility if you have time to make use of it |

Useful but time consuming

Dynamic

interesting, helpful, on course

Good tool - contacts - interesting - frustrating

easy access, varied content from trivia to useful, facilitating, good presentation, interesting.

Excellent facility. Extremly relevant to MBA course to use such technology.

Fun, stimulating

Interesting, same again

Interesting, helpful, facilitating

Great idea but represents a bit of a learning curve to the uninitiated.

assisting, interesting, helpful, stimulating.

Useful - particually with no local group.

good

useful but not a direct help for TMA's

different

Very interesting and helpful providing quality info in a related, informative way.

Does not add much value

ok

wordy, useful, stimulating

good discussion forum

time consuming

not very impressed

thread killer, obscuring, ineffective, unhelpful

First Class!!

good, facilitating, connecting.

helpful, stimulating

excellent

useful but could be improved with better direction in 'free' conferences

stimulating

ineffective, unhelpful

competitive advantage

The quality of FC is determined by the quality of contributions. Overall opinion is 'ok'

| | |
|---|---|
| fairly interesting | good. Conferencing adds a lot and e-mail chate & directory help you keep in touch with other students. |
| helpful, assisting, stimulating | |
| ambivalent | OK, adequate, developing, has things to learn. |
| generally helpful, assisting, effective | useful, difficult, treacle, time-consuming, community |
| good | Seems to have helped kill off self-help groups! |
| waste of time | good media, but I do wish I did not see a lot of irrelevant discussions |
| saitsfactory | |
| ineffective, showing off. | not particually well integrated. Not sure of its purpose. |
| Useful to ask specific questions. Conferences usually just a few keen contributors - moderators make a lot of effort to keep discussions going and come across as being very willing to help. | Informative educational facility |
| | Useful motivator + interesting place for discussions |
| unhelpful, irrelevant, time-intensive. | helpful, interesting, accessible |

50) Please use this space, and continue on the back if necessary, to make any additional comments about any aspect of the Active Archive.

---

[50-1] Its a tool! A dictionary is a tool. I have no interest in what you call either of them. Neither do I find either patronising, stimulating etc. I don't study either on a regular basis, nor do I want to "write down a few words " to describe them. I don't mind providing feedback but this level of detail (especially section 3) is rediculous.

[50-2] Very useful concept but would be better explained - i.e. wasn't sure what it was for a while then got the hang of it.

[50-3] Generally, I think it is good and an asset to my learning + development. It is a massive time saver - I would not look at the archive at all, even if it were not there, and might impoverish my learning as a result.

[50-4] As an active user of FC I have more or less ignored Active Archive - and still do not know what it is trying to acheive, so perhapse you should ignore my remarks. Though the fact that an active user still does not know might be significant.

[50-5] The active archive should probably only be used to stimulate discussion when or if it is flagging.

[50-6] Messages could be checked for grammer and spelling

[50-7] the active archive has been the most interesting part of first class for B882. Otherwise the conferences have been quite dull.

# Appendix 13: Notes from informal interviews with tutors.

This is a write up of my notes of the informal interviews I held with two of the computer tutors. Both interviews were quite informal, with the second one constrained to less than 10 minutes. I didn't start out with any great structure as to the questions I wanted to ask but rather settled on three areas. Because of the informal nature of the interviews there is a certain amount of overlap in their answers.

The first area I wanted to explore with them is what they thought of the concept of 'The Virtual Participant', moving on to how to make it better and more useable. I then wanted the tutors to reflect on the reality of how it had performed. In reality these two became tangled up in the responses. Finally I asked what their tutor group thought of it all.

## Tutor 1

With this tutor we had 20 minutes over coffee befor other meetings took place and we had to leave. I had also talked previously to this tutor during the training day for residential school tutors and I will note some of that discussion as well. The first thing they said was "if only it was true AI", where the response has always been that if the system was that good it would be worth the first Nobel prize in AI! The tutor was of course refuring to the recognition problem that we have with the system, identifying the right threads to link its knowledge base to. The next comment they made was that for the system to be of maximum use then the tutors should help with the structure. Their motivation for this was that a number of the AA responses are dated as the course has moved on, and also the types of student have changed. At the moment the tutor ignores it, but the concept is good. Having motivated experts, I believe, is half the battle for the acceptance of this system, the other half is an editing an control environment, something currently lacking. So my response to this was to ask 'how do we get tutor buy in? Or at least contributing to the up keep of the system?'. The answer to this was to incorporate it into the discussion, have the tutors follow up on the points it makes, and then feed this back into the database. Also only use the last two years worth of discussions (the VP has used discussions from 94/95/96 in both years) to keep the stuff current. It has been very useful on occasion as a trigger reminding of past discussions that might be relevant now.

We debated the use of the VP on other courses and its application to a more prescriptive course, perhapse on the managment certificiate or diploma programs where the course

has a core curriculum teaching well defined managment theory (the tutor went on to recomend a CD rom from the Institute of Managment as a source of this information). Certainly in a situation where the course schedual is more strictly controled, or case-studies are discussed then the VP could be more tightly focused. The tutor felt that the thing with B882 is it is a very rich course where the students are activly encouraged to go beyond the exsisting confines, however this does not mean that something totally irrelevant from the VP won't provide some help to someone (a sentiment reflected in the questionaire feedback).

I then asked about the reality of the system, what they thought about what had happend. To summarise some reflections 'its had its ups and downs but its certainly been interesting'. We then went on to discuss what would be the ideal system for the tutor which was a "virtual facilitator" something that could help by identfying discussions and presenting options to explore.

Prior to this interview I had spoken to this same tutor during the residential school tutor training day held about midway through the first block, a few weeks after the VP had started. At that point the tutor was less enthusiastic about the VP having been put off a little on its first appearance in the conference. The tutor reiterated the feeling that a lot of the stuff was too dated to be of particular use to the students. They also felt that the messages really only provided reflections on the current discussion. They again proposed, as others had done previously, that the messages should be posted to a seperate 'reflections' conference. I argued then, as I have befor, and do now, that this would completly negate any useful role of the system as the messages would not be contextualised to the current discussion. This idea was not mentioned again during the interview, and they appeared that they were quite happy with current approach. At this time the tutor again raised their fear that it might put students off from contributing, or even using conferencing.

## Tutor 2

I arranged to talk to this tutor at a time when they were visiting the university. Unfortunatly other things took over and I only got a chance to talk to them in the car driving them to the station. The tutor felt that overall they now understood and quite liked the concept. Looking back to when it was first introduced they did not feel that it had, or would, help them or save them time. Because system does not provide answers, something which is not really possible with the course I have been working with, the tutor felt the effect was to raise more ideas and points of discussion, and hence generate more work. In reality they felt it was easy on their part to ignore the

messages posted by the AA unless points were raised by other students. This meant that only those points that the students thought were important would be addressed by the tutors. Undoutably this is a reflection of the tutors lack of input, and hence commitment, to the system. During the course of the second study this tutor had to address points arising from the AA messages for two separate students. For this tutor the result of the second study was that they were willing to allow the system to be trialed on another set of students. In addition they were willing to explore other ways of using and improving the system.